

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 14, 2002, 04:24:33 ; Search time 75.1 seconds
(without alignments)
6308.815 Million cell updates/sec

Title: US-09-646-825-1

Perfect score: 2092

Sequence: 1 gaattctctagactccacca.....cttggttaagagctcaagctt 2092

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 351203 seqs, 113238999 residues

Total number of hits satisfying chosen parameters: 702406

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued_Patents_NA: *
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq: *
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq: *
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq: *
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq: *
5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq: *
6: /cgn2_6/ptodata/1/ina/backfiles1.seq: *

-Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query	Score	Match	Length	DB	ID	Description
1	Q1	147.6	7.1	694	4	US-08-998-416-741	Sequence 741, Appl
2		90.2	4.3	711	4	US-08-998-416-633	Sequence 633, Appl
3		48.6	2.3	7218	1	US-08-232-463-14	Sequence 14, Appl
4		39.2	1.9	7218	1	US-08-232-463-14	Sequence 14, Appl
5		35.6	1.7	964	1	US-07-798-223A-1	Sequence 1, Appl
6		34	1.6	4724	1	US-08-404-665-3	Sequence 3, Appl
7		34	1.6	4724	1	US-08-404-671-3	Sequence 3, Appl
8		34	1.6	4724	1	US-08-404-781-3	Sequence 3, Appl
9		33.6	1.6	4450	3	US-08-617-860B-2	Sequence 3, Appl
10		33.6	1.6	4706	2	US-08-331-081B-5	Sequence 5, Appl
11		33.6	1.6	4744	2	US-08-331-081B-3	Sequence 3, Appl
12		32.8	1.6	1293	1	US-08-476-008-43	Sequence 43, Appl
13		32.8	1.6	1293	1	US-08-306-063-43	Sequence 43, Appl
14		32.8	1.6	1293	1	US-08-833-485-43	Sequence 43, Appl
15		32.8	1.6	1293	4	US-09-137-440-43	Sequence 43, Appl
16		32	1.5	5191	1	US-08-340-428B-1	Sequence 1, Appl
17		32	1.5	5191	5	PCT-US93-07306-1	Sequence 1, Appl
18		31.8	1.5	1508	3	US-09-041-991A-3	Sequence 3, Appl
19		31.6	1.5	246240	2	US-08-724-394A-20	Sequence 20, Appl
20		31.6	1.5	246240	2	US-08-724-394A-21	Sequence 21, Appl
21		31.6	1.5	246240	2	US-08-724-394A-22	Sequence 22, Appl
22		31	1.5	432	4	US-09-008-697A-19	Sequence 19, Appl
23		30.6	1.5	289	4	US-09-007-005-17	Sequence 17, Appl
24		30.6	1.5	289	4	US-09-244-796-17	Sequence 17, Appl
25		30.6	1.5	3656	1	US-08-232-463-8	Sequence 8, Appl
26		30.6	1.5	9454	1	US-08-232-463-3	Sequence 3, Appl
27		30.6	1.5	9454	1	US-08-232-463-4	Sequence 4, Appl

28 30.2 1.4 573 1 US-08-102-757-20 Sequence 20, Appl
29 30.2 1.4 1645 5 PCT-US94-01321-9 Sequence 9, Appl
30 30 1.4 1083 2 US-09-057-762-24 Sequence 24, Appl
31 30 1.4 1083 3 US-08-326-119A-24 Sequence 24, Appl
32 30 1.4 2364 2 US-08-838-219B-5 Sequence 5, Appl
33 30 1.4 2364 3 US-09-233-336A-5 Sequence 5, Appl
34 30 1.4 2364 3 US-09-233-752A-5 Sequence 5, Appl
35 30 1.4 2364 4 US-09-402-036-5 Sequence 5, Appl
36 30 1.4 2375 4 US-08-960-780-3 Sequence 3, Appl
37 30 1.4 2375 4 US-08-960-780-7 Sequence 7, Appl
38 30 1.4 2375 4 US-09-073-898-3 Sequence 3, Appl
39 30 1.4 2375 4 US-09-073-898-7 Sequence 3, Appl
40 30 1.4 2378 1 US-08-471-033-28 Sequence 28, Appl
41 30 1.4 2378 2 US-08-471-044-28 Sequence 28, Appl
42 30 1.4 2378 2 US-08-463-483A-28 Sequence 28, Appl
43 30 1.4 2378 2 US-08-471-046A-28 Sequence 28, Appl
44 30 1.4 2378 2 US-08-470-566B-28 Sequence 28, Appl
45 30 1.4 2378 2 US-08-838-219B-1 Sequence 1, Appl

ALIGNMENTS

RESULT 1
US-08-998-416-741
; Sequence 741, Application US/08998416
; Patent No. 6239264
; GENERAL INFORMATION:
; APPLICANT: Philippsen, Peter
; APPLICANT: Pohlmann, Rainer
; APPLICANT: Steiner, Sabine
; APPLICANT: Mohr, Christine
; APPLICANT: Wendland, Jurgen
; APPLICANT: Knechtie, Philipp
; APPLICANT: Rebeschung, Corinne
; TITLE OF INVENTION: GENOMIC DNA SEQUENCES OF ASHBYA GOSSYPII
; NUMBER OF SEQUENCES: 1152
; CORRESPONDENCE ADDRESS:
; ADDRESSES: No. 6239264artis Corporation
; STREET: 3054 Cornwallis Road
; CITY: Research Triangle Park
; STATE: No. 6239264th Carolina
; COUNTRY: USA
; ZIP: 27709
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/998,416
; FILING DATE: 24-DEC-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: CH 0016/97
; FILING DATE: 31-DEC-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Meigs, J. Timothy
; REGISTRATION NUMBER: 38,241
; REFERENCE/DOCKET NUMBER: PF/5-30306/A/CGC1976
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919-541-8587
; TELEFAX: 919-541-8689
; INFORMATION FOR SEQ ID NO: 741:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 694 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; ORIGINAL SOURCE:
; ORGANISM: PAG1482RP

US-08-998-416-741

Query Match 7.1% Score 147.6; DB 4; Length 694;
Best Local Similarity 54.68; Pred. No. 1.6e-38;
Matches 344; Conservative 0; Mismatches 274; Indels 12; Gaps 2;

Qy 1006 ctacttcagatgggaatagtagcacaaattcttatgtccatcatcatctttccagtcgca 1065
Db 76 CTATTTTAAATGGGGTGTGTGCAACGGTTTGGCCGGGCTTCTGCTCGCATATGA 135
Qy 1066 gaaggtcttaaggaacagaggttatgaaattctttactttatttcacaaagccatgaacat 1125
Db 136 AAAGTTTATTCGCCAAAGAACGTAACGAATCTTCTGTATCTGACACAGCTCTTCAACAT 195
Qy 1126 catgtttatcatagctatgtattaccattgccacacactagagtgagtgagctgactg 1185
Db 196 TGTCTTTTATGTCATGTATATGCAATCAAAACGCTGGGATGGCAGGCTGGGTCTG 255
Qy 1186 gtccatggctggcctctcttcagcaggttctgcgaattgtacgtatcatcatgaa 1245
Db 256 GTCGATGTTGCCATCTACTCTTCGAGCGTGTGCCCGGATAGCTCGATTGTACTTGC 315
Qy 1246 cggaggtcttaagaccgccactttgtcgacacagatgattcttaacgtttacagatctc 1305
Db 316 TGGAGGCATCAAGAGGCCACATT---AACAGATGTTGGGGATCGCGTGTCTCAAGATGAC 372
Qy 1306 tgtcaagaagcctaagttcttcagatcatcaagtggagcatttgccatatgtactttct 1365
Db 373 AGTGAGAGCCAAAGCATTTCAATATATACCCGGGGCTTATGTTTTCGTTTATTTAT 432
Qy 1366 ttccacaaatcagcgtgtgtctacagttttcaatctcatccttcacagtcctcatcaga 1425
Db 433 TAGTGGGAAGGATGCTTGTCTTATCCATTCAGTCGCACCGTTCACCGTCTTAATAC 492
Qy 1426 aaggcacagagatccttaacacccagatcacaactaactatgtacgtcaaaagctaaaggg 1485
Db 493 ACCCAAGATCGATGGCACAACCTGGTG-----ATTTATTTCAAGTGCACAAGGG 543
Qy 1486 cattacagagatcctcttaagaaagtcttaagcgtccaaacacataccgttgattgcaa 1545
Db 544 CGTACGCGAGAGCTGTCTTAACAGAGATCTTCTATCCGGGAAGAGTCCATCGAATACA 603
Qy 1546 gattttcttagaggaacatagcgttaactgtccctcacattgccaacttaagagaaa 1605
Db 604 GGTGCTTCTAGAAAGGCCCTATGAAACACCATTCGCGGGCTTGTGCTCTGACCGCG 663
Qy 1606 tctagtaggtagctgcgggctcgcgt 1635
Db 664 CTACGTGGGCGCCAGCGAGGTCTTGGCGT 693

RESULT 2

US-08-998-416-633
; Sequence 633, Application US/08998416
; Patent No. 6239264
; GENERAL INFORMATION:
; APPLICANT: Philippsen, Peter
; APPLICANT: Pohlmann, Rainer
; APPLICANT: Steiner, Sabine
; APPLICANT: Mohr, Christine
; APPLICANT: Wendland, Jurgen
; APPLICANT: Knechtle, Philipp
; APPLICANT: Reibschung, Corinne
; TITLE OF INVENTION: GENOMIC DNA SEQUENCES OF ASHBYA GOSSYPII
; NUMBER OF INVENTION: AND USES THEREOF
; CORRESPONDENCE ADDRESSES:
; ADDRESS: No. 6239264artis Corporation
; STREET: 3054 Cornwallis Road
; CITY: Research Triangle Park
; STATE: No. 6239264th Carolina
; COUNTRY: USA

ZIP: 27709
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: IBM PC compatible
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/998,416
FILING DATE: 24-DEC-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: CH 0016/97
FILING DATE: 31-DEC-1996
ATTORNEY/AGENT INFORMATION:
NAME: Meigs, J. Timothy
REGISTRATION NUMBER: 38,241
REFERENCE/DOCKET NUMBER: PF/5-30306/A/CGC1976
TELECOMMUNICATION INFORMATION:
TELEPHONE: 919-541-8587
TELEFAX: 919-541-8689
INFORMATION FOR SEQ ID NO: 633:
SEQUENCE CHARACTERISTICS:
LENGTH: 711 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
ORIGINAL SOURCE: PAG1427RP
ORGANISM: US-08-998-416-633

Query Match 4.3% Score 90.2; DB 4; Length 711;
Best Local Similarity 50.38; Pred. No. 1.4e-19;
Matches 280; Conservative 0; Mismatches 268; Indels 9; Gaps 2;
Qy 1526 aaccataccgttgattgcaagattttcttagaggaccatagggcgttaactgctccctcac 1585
Db 17 AAAGAGTCCATCGAATACAAAGTGCTTCTAGAAAGGCCCTATGGAACACCATTCGCGG 76
Qy 1586 attgccaaacttaagagaaatcttagtagtagtgcctgagcgttcgagcagcctc 1645
Db 77 CTGCTGCTCTGACCGCGGTACGTGGCGGCCAGCGAGGTCTTGGCGTACGACGCGTGC 136
Qy 1646 taccoccatctctagaatgcttagattgcttagcactgacactgcagcagcagttc 1705
Db 137 TACCACACTTCGT-----CTCTCTGTGGACAAAGAACCCAGTTTCAACCATTCATTC 190
Qy 1706 tactggatcgtaacgaccttagtcaaccttaagtgttcgaaacagagctacaattggtctt 1765
Db 191 TACTGGATTATAAATGACCTTTTCATATCTGCATTTGTTTTCGCATGAGCTGAGGTACCTG 250
Qy 1766 aaggagaaattcttgaaagtctctgtctacactggttcacatggttcagtgagagatacaac 1825
Db 251 GCGGACCGGAATCGGACATCAAAATTTATACAGAGGAGCAATGAGTCGGCTAAAGAA 310
Qy 1826 tcatgagtcctccactaagggtttcgt---atgacaaggagaatctgaaatcacctagaa 1882
Db 311 CTGACCCAGATGTGCGCATTCGCCCTCTCGGAAGTTCTGTTGGATTCTGCTGGATATCTGC 370
Qy 1883 tgccttaacagagggccagacctcaaaagagtagtagatcagagatcaaaattgtcagaa 1942
Db 371 AGGTCCTCTCTGCGCCAGATCTCAAGAGATCGTGGAAAGAGCAGTCTGCTCTCGTCT 430
Qy 1943 ctgagagaacaacacatcaactttctactcatcgagacagagcactttcaatgacgacttt 2002
Db 431 AACGAGGCACAGGAGCTCACGCTTTATTAGCAGCGGCCCTTCGACCTTTAATGACCATTTTC 490
Qy 2003 aggaatcaggttgataaaggtatgacttcttagctgaagatagatgtcgaaactagagag 2062
Db 491 CGCTATGCTGTGAATCTAGCATACGCGGCAAACTCCAGTGTGATGTCGACCTAGAGGAG 550
Qy 2063 gagagtttacttggtta 2079


```

; LENGTH: 4744 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-331-081B-3

```

		Query Match	1.6%	Score 33.6;	DB 2;	Length 4744;
		Best Local Similarity	55.0%;	Pred. No. 2,2;		
		Matches 66;	Conservative 0;	Mismatches 54;	Indels 0;	Gaps 0;
QY	364	agttgttccacaccggttgatgcgaatggagcggccctatcaactactactatgatggaaaaa	423			
Db	3448	AGTTATTACTAACATACCTTTCGAATGAGAGACACAAAAACAATGCACGTGAGAAA	3389			
QY	424	ctaaggatcacacttgaaattgatcgcatctcaatgfcgcatggggcccgctctcttt	483			
Db	3388	TACAGAAGCGCATCAAAATTTGAAGGCCAAGATGGTTCCGACCCTTGTCATGTGCTCTCTT	3329			

S
-08-476-008-43
Sequence 43, Application US/08476008
; Patent No. 5627061
; GENERAL INFORMATION:
; APPLICANT: Barry, Gerard F.
; APPLICANT: Kishore, Ganesh M.
; APPLICANT: Padgett, Stephen R.
; APPLICANT: Stallings, William C.
; TITLE OF INVENTION: Glycophosphate Tolerant
; TITLE OF INVENTION: 5-Enolpyruvylshikimate-3-Phosphate Synthases

```

; NAME/KEY: CDS
; LOCATION: 1..1293
; US-08-476-008-43

Query Match 1.6%; Score 32.8; DB 1; Length 1293;
Best Local Similarity 45.9%; Pred. No. 1.7; 132; Indels 0; Gaps
Matches 112; Conservative 0; Mismatches 132;

QY 370 ttcaacacggttgatggcaaatgagacgcgcctatcactactactactatgaggaactatgg 429
    ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 764 TTCATAATGTTGGAATCAATCAAAACACGTTTCAGGTATTATTGATATTCTTCAAAAAATGG 823

QY 430 gatccactgaattgatgcgatctcaatggtgogcatggyggccctgcctctctctctgggt 489
    ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 824 GCGGTAAATATCCAACCTTTTCAATCAAAACAACTGGTGCCTGAACTACTCTCTTCTATTTCGTA 883

QY 490 cgcagtcottaccgcgcgaactatcttgaacatctcctcaaacgcgtattcggcaagaacat 549
    ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 884 TTCAAATACACACAAATGCTTCAACCAATAAACAAATCGAAGGAGAATTAGTTCCTCAAAAGCAA 943

QY 550 tatggcaaatctgttaagaagtcctcttatctaccaccaagcgtttacaaaagactacaacga 609
    ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 944 TTGATGAACCTGCTGTAAATAGCATTAACCTTTGTACACAGCAGTTCGCACGAGTACAATTA 1003

QY 610 gaga 613
    ||||
Db 1004 AAGA 1007

RESULT 13
US-08-306-063-43
; Sequence 43, Application US/08306063
; Patent No. 5633435
; GENERAL INFORMATION:
; APPLICANT: Barry, Gerard F.
; APPLICANT: Kishore, Ganesh M.
; APPLICANT: Padgett, Stephen R.
; APPLICANT: Stallings, William C.
; TITLE OF INVENTION: Glyphosate Tolerant
; TITLE OF INVENTION: 5-Etholpyruvylshikimate-3-Phosphate Synthases
; NUMBER OF SEQUENCES: 69
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dennis R. Hoerner, Jr., Monsanto Co. BB4F
; STREET: 700 Chesterfield Village Parkway
; CITY: St. Louis
; STATE: Missouri
; COUNTRY: USA
; ZIP: 63198
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/306.063
; FILING DATE: 13-SEP-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/749,611
; FILING DATE: 28-AUG-1991
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; APPLICATION NUMBER: US 07/576,537
; FILING DATE: 31-AUG-1990
; CLASSIFICATION: 435
; NAME: Hoerner Jr., Dennis R.
; REGISTRATION NUMBER: 30,914
; REFERENCE/DOCKET NUMBER: 38-21(10660)A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (314)537-6099
; TELEFAX: (314)537-6047

```

INFORMATION FOR SEQ ID NO: 43:
SEQUENCE CHARACTERISTICS:
LENGTH: 1293 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 1..1293
US-08-306-063-43

Query Match 1.6% Score 32.8; DB 1; Length 1293;
Best Local Similarity 45.9%; Pred. No. 1.7; Indels 0; Gaps 0;
Matches 112; Conservative 0; Mismatches 132; Indels 0; Gaps 0;
QY 370 ttcaaacacggttgatggcaaatgagagcgccctatctactactatgaggaactatgg 429
DB 764 TTCATAATGTTGGAATCAATCAACACGTTTCAGGTATTATTGATATTGTTGAAAAATGG 823
QY 430 gatccactgaattgatgcatcgaatggtgcgcatggggcgctctctctctctgggt 489
DB 824 GCGTAATATCCAACTTTTCAATCAACAACTGGTGTGAACCTACTGCTTCTATTGTA 883
QY 490 cgcagtcctaccgcccgaactctctgaacattctcaaacgcgtatttcggcaagaacat 549
DB 884 TTCAATACACCAACTGCTTCAACCAATACAAATCGAAGGAGAAATAGTTCCAAAGCAA 943
QY 550 tatggcaaatctgttaagaagtcctctatctaccacgaagctttacaagaactacaaga 609
DB 944 TTGATGAACCTGCCCTGTAATAGCATTACTTTGTACACAAAGCAGTTGGCAGCAGTACAATTA 1003
QY 610 gaga 613
DB 1004 AAGA 1007

RESULT 14

US-08-833-485-43
Sequence 43, Application US/08833485
Patent No. 5804425
GENERAL INFORMATION:
APPLICANT: Barry, Gerard F.
APPLICANT: Kishore, Ganesh M.
APPLICANT: Padgett, Stephen R.
APPLICANT: Stallings, William C.
TITLE OF INVENTION: Glyphosate Tolerant
TITLE OF INVENTION: 5-Enolpyruvylshikimate-3-Phosphate Synthases
NUMBER OF SEQUENCES: 69
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dennis R. Hoerner, Jr., Monsanto Co. BB4F
STREET: 700 Chesterfield Village Parkway
CITY: St. Louis
STATE: Missouri
COUNTRY: USA
ZIP: 63198
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/833,485
FILING DATE: 07-APR-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/306,063
FILING DATE: 13-SEP-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/749,611
FILING DATE: 28-AUG-1991

CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/576,537
FILING DATE: 31-AUG-1990
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Hoerner Jr., Dennis R.
REGISTRATION NUMBER: 30,914
REFERENCE/DOCKET NUMBER: 38-21(15117)A
TELEPHONE: (314)737-6099
TELEFAX: (314)737-6047
INFORMATION FOR SEQ ID NO: 43:
SEQUENCE CHARACTERISTICS:
LENGTH: 1293 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 1..1293
US-08-833-485-43

Query Match 1.6% Score 32.8; DB 1; Length 1293;
Best Local Similarity 45.9%; Pred. No. 1.7; Indels 0; Gaps 0;
Matches 112; Conservative 0; Mismatches 132; Indels 0; Gaps 0;
QY 370 ttcaaacacggttgatggcaaatgagagcgccctatctactactatgaggaactatgg 429
DB 764 TTCATAATGTTGGAATCAATCAACACGTTTCAGGTATTATTGATATTGTTGAAAAATGG 823
QY 430 gatccactgaattgatgcatcgaatggtgcgcatggggcgctctctctctctgggt 489
DB 824 GCGTAATATCCAACTTTTCAATCAACAACTGGTGTGAACCTACTGCTTCTATTGTA 883
QY 490 cgcagtcctaccgcccgaactctctgaacattctcaaacgcgtatttcggcaagaacat 549
DB 884 TTCAATACACCAACTGCTTCAACCAATACAAATCGAAGGAGAAATAGTTCCAAAGCAA 943
QY 550 tatggcaaatctgttaagaagtcctctatctaccacgaagctttacaagaactacaaga 609
DB 944 TTGATGAACCTGCCCTGTAATAGCATTACTTTGTACACAAAGCAGTTGGCAGCAGTACAATTA 1003
QY 610 gaga 613
DB 1004 AAGA 1007

RESULT 15

US-09-137-440-43
Sequence 43, Application US/09137440
Patent No. 6248876
GENERAL INFORMATION:
APPLICANT: Barry, Gerard F.
APPLICANT: Kishore, Ganesh M.
APPLICANT: Padgett, Stephen R.
APPLICANT: Stallings, William C.
TITLE OF INVENTION: Glyphosate Tolerant
TITLE OF INVENTION: 5-Enolpyruvylshikimate-3-Phosphate Synthases
NUMBER OF SEQUENCES: 69
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dennis R. Hoerner, Jr., Monsanto Co. BB4F
STREET: 700 Chesterfield Village Parkway
CITY: St. Louis
STATE: Missouri
COUNTRY: USA
ZIP: 63198
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS

```

; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/137,440
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/833,485
; FILING DATE: 07-APR-1997
; APPLICATION NUMBER: US 08/306,063
; FILING DATE: 13-SEP-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/749,611
; FILING DATE: 28-AUG-1991
; APPLICATION DATA:
; APPLICATION NUMBER: US 07/576,537
; FILING DATE: 31-AUG-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Hoerner Jr., Dennis R.
; REGISTRATION NUMBER: 30,914
; REFERENCE/DOCKET NUMBER: 38-21(15117)A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (314)737-6099
; TELEFAX: (314)737-6047
; INFORMATION FOR SEQ ID NO: 43:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1293 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..1293
; US-09-137-440-43

```

Query Match	1.6%	Score 32.8;	DB 4;	Length 1293;
Best Local Similarity	45.9%	Pred. No. 1.7;		
Matches 112;	Conservative 0;	Mismatches 132;	Indels 0;	Gaps 0;
Qy	370	ttcecaaccgttgatggcaaatgagacggccctatcactactactatgagggaaaactatgg	429	
Db	764	TTcATAATGTTGGAAATCAATCAAAACACGTCAGGTATTATTGATATTGTTGAAAAATGG	823	
Qy	430	gatccacttgaattgatgcgatccaatggtgcgcgatggggcctcgctctctctctggt	489	
Db	824	CGGTAATATCCAACTTTTCAATCAAAACACTGGTGTGGAACCTACTGCTTCTATTTCGTA	883	
Qy	490	cgcagctctaccgccgcaactatcttgaacattctcaaacgcgtatttcggcaagaacat	549	
Db	884	TTCAATACACACCAATGCTTCAACCAATACAAATCGAAGGAGAAATTAGTTCCAAAGCAA	943	
Qy	550	tatggcaaatcttgttaagatctcttatctaccacagcgttttacaagaactacaaga	609	
Db	944	TTGATGAACTCGCCGTAAATAGCATTTACTTTGTACACAAGCAGTTGGCACGAGTACAAATTA	1003	
Qy	610	gaga	613	
Db	1004	AAGA	1007	

Search completed: May 14, 2002, 07:25:16
Job time: 10843 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 14, 2002, 05:52:23 ; Search time 186.08 Seconds
(without alignments)
9638.459 Million cell updates/sec

Title: US-09-646-825-1

Perfect score: 2092

Sequence: 1 gaattcttagactccacca.....cttggttaagagctcaagctt 2092

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 930621 seqs, 42862619 residues

Total number of hits satisfying chosen parameters: 1861242

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N_Geneseq_1101.*
1: /SIDSI/gcgdata/geneseq/geneseq/NA1980.DAT.*
2: /SIDSI/gcgdata/geneseq/geneseq/NA1981.DAT.*
3: /SIDSI/gcgdata/geneseq/geneseq/NA1982.DAT.*
4: /SIDSI/gcgdata/geneseq/geneseq/NA1983.DAT.*
5: /SIDSI/gcgdata/geneseq/geneseq/NA1984.DAT.*
6: /SIDSI/gcgdata/geneseq/geneseq/NA1985.DAT.*
7: /SIDSI/gcgdata/geneseq/geneseq/NA1986.DAT.*
8: /SIDSI/gcgdata/geneseq/geneseq/NA1987.DAT.*
9: /SIDSI/gcgdata/geneseq/geneseq/NA1988.DAT.*
10: /SIDSI/gcgdata/geneseq/geneseq/NA1989.DAT.*
11: /SIDSI/gcgdata/geneseq/geneseq/NA1990.DAT.*
12: /SIDSI/gcgdata/geneseq/geneseq/NA1991.DAT.*
13: /SIDSI/gcgdata/geneseq/geneseq/NA1992.DAT.*
14: /SIDSI/gcgdata/geneseq/geneseq/NA1993.DAT.*
15: /SIDSI/gcgdata/geneseq/geneseq/NA1994.DAT.*
16: /SIDSI/gcgdata/geneseq/geneseq/NA1995.DAT.*
17: /SIDSI/gcgdata/geneseq/geneseq/NA1996.DAT.*
18: /SIDSI/gcgdata/geneseq/geneseq/NA1997.DAT.*
19: /SIDSI/gcgdata/geneseq/geneseq/NA1998.DAT.*
20: /SIDSI/gcgdata/geneseq/geneseq/NA1999.DAT.*
21: /SIDSI/gcgdata/geneseq/geneseq/NA2000.DAT.*
22: /SIDSI/gcgdata/geneseq/geneseq/NA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query	Score	Match	Length	DB	ID	Description
1	2092	100.0	2092	20	AAZ59028		Codon optimised ye
2	1239.8	59.3	2059	20	AAZ59061		Wild type yeast Fe
3	83	4.0	83	20	AAZ59031		Oligonucleotide A-
4	83	4.0	83	20	AAZ59032		Oligonucleotide A-
5	83	4.0	83	20	AAZ59033		Oligonucleotide A-
6	83	4.0	83	20	AAZ59034		Oligonucleotide A-
7	83	4.0	83	20	AAZ59035		Oligonucleotide A-
8	83	4.0	83	20	AAZ59036		Oligonucleotide A-
9	83	4.0	83	20	AAZ59043		Oligonucleotide C-
10	83	4.0	83	20	AAZ59044		Oligonucleotide C-
11	83	4.0	83	20	AAZ59045		Oligonucleotide C-

C 12	83	4.0	83	20	AAZ59046	Oligonucleotide C-
C 13	83	4.0	83	20	AAZ59047	Oligonucleotide C-
C 14	83	4.0	83	20	AAZ59048	Oligonucleotide C-
C 15	82	3.9	82	20	AAZ59049	Oligonucleotide D-
C 16	82	3.9	82	20	AAZ59050	Oligonucleotide D-
C 17	82	3.9	82	20	AAZ59051	Oligonucleotide D-
C 18	82	3.9	82	20	AAZ59052	Oligonucleotide D-
C 19	82	3.9	82	20	AAZ59053	Oligonucleotide D-
C 20	82	3.9	82	20	AAZ59054	Oligonucleotide D-
C 21	80.6	3.9	936	22	AAF58252	Oligonucleotide D1
C 22	80.6	3.9	936	22	AAF58252	Oligonucleotide D1
C 23	80.6	3.9	936	22	AAF58257	Oligonucleotide D1
C 24	80.6	3.9	936	22	AAF58259	Oligonucleotide D2
C 25	80.6	3.9	936	22	AAF58262	Oligonucleotide D2
C 26	80.6	3.9	938	22	AAF58255	Oligonucleotide D1
C 27	80	3.8	80	20	AAZ59037	Oligonucleotide B-
C 28	80	3.8	80	20	AAZ59038	Oligonucleotide B-
C 29	80	3.8	80	20	AAZ59039	Oligonucleotide B-
C 30	80	3.8	80	20	AAZ59040	Oligonucleotide B-
C 31	80	3.8	80	20	AAZ59041	Oligonucleotide B-
C 32	80	3.8	80	20	AAZ59042	Oligonucleotide B-
C 33	77	3.7	77	20	AAZ59055	Oligonucleotide E-
C 34	77	3.7	77	20	AAZ59056	Oligonucleotide E-
C 35	77	3.7	77	20	AAZ59057	Oligonucleotide E-
C 36	77	3.7	77	20	AAZ59058	Oligonucleotide E-
C 37	77	3.7	77	20	AAZ59059	Oligonucleotide E-
C 38	77	3.7	77	20	AAZ59060	Oligonucleotide E-
C 39	75.6	3.6	936	22	AAF58252	Oligonucleotide D1
C 40	75.6	3.6	936	22	AAF58254	Oligonucleotide D1
C 41	75.6	3.6	936	22	AAF58257	Oligonucleotide D1
C 42	75.6	3.6	936	22	AAF58259	Oligonucleotide D2
C 43	75.6	3.6	936	22	AAF58262	Oligonucleotide D2
C 44	75.6	3.6	938	22	AAF58255	Oligonucleotide D1
C 45	43	2.1	4590	22	AAH24065	Yeast AOD9604-asso

ALIGNMENTS

RESULT 1

AAZ59028
ID AAZ59028 standard; DNA; 2092 BP.
XX
AC AAZ59028;
XX
DT 11-APR-2000 (first entry)
XX
DE Codon optimised yeast Fe3+-reductase gene.
XX
KW Crop plant; polyadenylation signal; substitution; ferric reductase; iron;
KW yeast; heterologous gene; ds.
XX
OS Saccharomyces cerevisiae.
OS Synthetic.
XX
FH Key Location/Qualifiers
FT CDS 20..2080
FT /*tag= a
FT /product= "Fe3+-reductase"
XX
PN W09948356-A1.
XX
PD 30-SEP-1999.
XX
PF 24-MAR-1999; 99WO-JP01481.
XX
PR 24-MAR-1998; 98JP-0096637.
XX
PA (NISC-) JAPAN SCI & TECHNOLOGY CORP.
XX
PI Mori S, Nakanishi H, Oki H, Yamaguchi H;
XX
DR WPI; 1999-571924/48.

DR P-PSDB: AAY56088.

... PT Transformation of higher plant with foreign gene having modified
PT poly(A) site -
PT

XX
F047(11)

PS Claim 20; Fig 9; 81pp; Japanese.

The invention relates to a method of transforming a higher plant with a foreign gene in which the polyadenylation signal region contained in the nucleotide sequence of the foreign gene has been modified by nucleotide substitution to give a sequence not having this function, in such a way that the function of the foreign gene expression product is not substantially impaired. The foreign gene used is especially an iron (Fe) 3+ (ferric) reductase gene in which putative polyadenylation signals contained within the coding sequence are substituted with alternative codons for expression in higher plants. This sequence represents a plant codon optimised version of the wild type yeast ferric reductase (FRR1 gene). The sequence has had putative polyadenylation signal and poly(A) addition sites removed by nucleotide substitution. The method allows the production of transformant plants with improved expression of a desired foreign gene, especially to improve iron absorption efficiency in crop plants.

SQ Sequence 2092 BP; 587 A; 510 C; 418 G; 577 T; 0 other;

Query Match	Score 2092;	DB 20;	Length 2092;
Best Local Similarity	100.0%;		
Matches 2092;	Conservative	0;	Mismatches
Indels	0;	Gaps	0;

Qy	1	gaattctctagactccaacattggttagaacaacagagtcctttctgctcttcaatctcttt	60
Db	1	gaattctctagactccaacattggttagaacaacagagtcctttctgctcttcaatctcttt	60
Qy	61	cttcgtcacagtcacactcgagcgtacacactcaactcctcaactctgcatcttcaggctgc	120
Db	61	cttcgtcacagtcacactcgagcgtacacactcaactcctcaactctgcatcttcaggctgc	120
Qy	121	actgtacacagttcggatgctcaagcaagtcacaaagtcttgctactgcaagaacatcaattg	180
Db	121	actgtacacagttcggatgctcaagcaagtcacaaagtcttgctactgcaagaacatcaattg	180
Qy	181	gctcggaaagcgtcactgcattgcattgagaactcaaatctcaacaaagctctgcgactc	240
Db	181	gctcggaaagcgtcactgcattgcattgagaactcaaatctcaacaaagctctgcgactc	240
Qy	241	cgctttgatgaaccttgccagccaatgctcaagtatacaagttttcacactcgaggacat	300
Db	241	cgctttgatgaaccttgccagccaatgctcaagtatacaagttttcacactcgaggacat	300
Qy	301	gaagaacaactacatttaatgcaagtaactaccttcgcgtctcttgagaaatccgataagaa	360
Db	301	gaagaacaactacatttaatgcaagtaactaccttcgcgtctcttgagaaatccgataagaa	360
Qy	361	gacagttgtttcacaaacgcttgatggcaaatgagcggcctatacactactactatgaga	420
Db	361	gacagttgtttcacaaacgcttgatggcaaatgagcggcctatacactactactatgaga	420
Qy	421	aaactatgggataccacttgattgtagcgtatcgaatggtcgcatgggggctctctt	480
Db	421	aaactatgggataccacttgattgtagcgtatcgaatggtcgcatgggggctctctt	480
Qy	481	cttctgggtcgcagtccttaaccgcgcacatacttgacaatctctcaaacgcgtattcgg	540
Db	481	cttctgggtcgcagtccttaaccgcgcacatacttgacaatctctcaaacgcgtattcgg	540
Qy	541	caagaacatattgcaaaattctgttaagaagtcctttatctaccgaagcgtttcacaaaga	600
Db	541	caagaacatattgcaaaattctgttaagaagtcctttatctaccgaagcgtttcacaaaga	600
Qy	601	ctacaacagagaactttctattctttggaacggttgccattccaactttacaactcgagg	660
Db	601	ctacaacagagaactttctattctttggaacggttgccattccaactttacaactcgagg	660

DE AA259032 standard; DNA; 83 BP.
XX AC AA259032;
XX DT 11-APR-2000 (first entry)
XX DE Oligonucleotide A-2 for generating codon-optimised ferric reductase.
XX KW Crop plant; polyadenylation signal; substitution; ferric reductase; iron;
XX KW yeast; heterologous gene; ss.
XX OS Saccharomyces cerevisiae.
XX OS Synthetic.
XX PN WO9948356-A1.
XX PD 30-SEP-1999.
XX PF 24-MAR-1999; 99WO-JP01481.
XX PR 24-MAR-1998; 98JP-0096637.
XX PA (NISC-) JAPAN SCI & TECHNOLOGY CORP.
XX PI Mori S, Nakanishi H, Oki H, Yamaguchi H;
XX DR WPI; 1999-571924/48.
XX PT Transformation of higher plant with foreign gene having modified
XX PT poly(A) site -
XX PS Disclosure; Fig 6; 81pp; Japanese.
XX CC The invention relates to a method of transforming a higher plant with
a foreign gene in which the polyadenylation signal region contained in
the nucleotide sequence of the foreign gene has been modified by
nucleotide substitution to give a sequence not having this function, in
such a way that the function of the foreign gene expression product is
not substantially impaired. The foreign gene used is especially an iron
(Fe) 3+ (ferric) reductase gene in which putative polyadenylation
signals contained within the coding sequence are substituted with
alternative codons for expression in higher plants. Oligonucleotides
AA259031-259060 were used as overlapping oligonucleotide primers to
generate a plant codon-optimised yeast ferric reductase (FREL) gene;
AA259028) in which putative polyadenylation signal and poly(A) addition
sites are removed by nucleotide substitution. The method allows the
production of transformant plants with improved expression of a desired
foreign gene, especially to improve iron absorption efficiency in crop
plants.
SQ Sequence 83 BP; 21 A; 26 C; 15 G; 21 T; 0 other;
Query Match 4.0%; Score 83; DB 20; Length 83;
Best Local Similarity 100.0%; Pred. No. 5.2e-16;
Matches 83; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 71 gtccaatcagcgctacatcctcattcgcattcatttcaggctgacctgaccag 130
Db 1 gtccaatcagcgctacatcctcattcgcattcatttcaggctgacctgaccag 60
QY 131 ttccggtatgcagcaagtcacaa 153
Db 61 ttccggtatgcagcaagtcacaa 83
RESULT 5
ID AA259033
XX AA259033 standard; DNA; 83 BP.
XX AC AA259033;
XX DT 11-APR-2000 (first entry)

XX DE Oligonucleotide A-3 for generating codon-optimised ferric reductase.
XX AC Crop plant; polyadenylation signal; substitution; ferric reductase; iron;
XX KW yeast; heterologous gene; ss.
XX OS Saccharomyces cerevisiae.
XX OS Synthetic.
XX PN WO9948356-A1.
XX PD 30-SEP-1999.
XX PF 24-MAR-1999; 99WO-JP01481.
XX PR 24-MAR-1998; 98JP-0096637.
XX PA (NISC-) JAPAN SCI & TECHNOLOGY CORP.
XX PI Mori S, Nakanishi H, Oki H, Yamaguchi H;
XX DR WPI; 1999-571924/48.
XX PT Transformation of higher plant with foreign gene having modified
XX PT poly(A) site -
XX PS Disclosure; Fig 6; 81pp; Japanese.
XX CC The invention relates to a method of transforming a higher plant with
a foreign gene in which the polyadenylation signal region contained in
the nucleotide sequence of the foreign gene has been modified by
nucleotide substitution to give a sequence not having this function, in
such a way that the function of the foreign gene expression product is
not substantially impaired. The foreign gene used is especially an iron
(Fe) 3+ (ferric) reductase gene in which putative polyadenylation
signals contained within the coding sequence are substituted with
alternative codons for expression in higher plants. Oligonucleotides
AA259031-259060 were used as overlapping oligonucleotide primers to
generate a plant codon-optimised yeast ferric reductase (FREL) gene;
AA259028) in which putative polyadenylation signal and poly(A) addition
sites are removed by nucleotide substitution. The method allows the
production of transformant plants with improved expression of a desired
foreign gene, especially to improve iron absorption efficiency in crop
plants.
SQ Sequence 83 BP; 26 A; 21 C; 17 G; 19 T; 0 other;
Query Match 4.0%; Score 83; DB 20; Length 83;
Best Local Similarity 100.0%; Pred. No. 5.2e-16;
Matches 83; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 141 caagcaagtcacaaagcttctgctactgcaagaacatcatcattgctcggaagcgctacatgcat 200
Db 1 caagcaagtcacaaagcttctgctactgcaagaacatcatcattgctcggaagcgctacatgcat 60
QY 201 gcgcttatgagaactcacaatct 223
Db 61 gcgcttatgagaactcacaatct 83
RESULT 6
ID AA259034/C
XX AA259034 standard; DNA; 83 BP.
XX AC AA259034;
XX DT 11-APR-2000 (first entry)
XX DE Oligonucleotide A-4 for generating codon-optimised ferric reductase.
XX KW Crop plant; polyadenylation signal; substitution; ferric reductase; iron;
XX KW yeast; heterologous gene; ss.

```

XX OS Saccharomyces cerevisiae.
XX OS Synthetic.
XX PN WO9948356-A1.
XX PD
XX PF
XX PI
XX PR
XX PP
XX PT
XX PS
XX CC The invention relates to a method of transforming a higher plant with
XX CC a foreign gene in which the polyadenylation signal region contained in
XX CC the nucleotide sequence of the foreign gene has been modified by
XX CC nucleotide substitution to give a sequence not having this function, in
XX CC such a way that the function of the foreign gene expression product is
XX CC not substantially impaired. The foreign gene used is especially an iron
XX CC (Fe) 3+ (ferric) reductase gene in which putative polyadenylation
XX CC signals contained within the coding sequence are substituted with
XX CC alternative codons for expression in higher plants. Oligonucleotides
XX CC AAZ59031-259060 were used as overlapping oligonucleotide primers to
XX CC generate a plant codon-optimised yeast ferric reductase (FREL gene;
XX CC AAZ59028) in which putative polyadenylation signal and poly(A) addition
XX CC sites are removed by nucleotide substitution. The method allows the
XX CC production of transformant plants with improved expression of a desired
XX CC foreign gene, especially to improve iron absorption efficiency in crop
XX CC plants.
XX SQ Sequence 83 BP; 20 A; 14 C; 22 G; 27 T; 0 other;

Query Match 4.0%; Score 83; DB 20; Length 83;
Best Local Similarity 100.0%; Pred. No. 5.2e-16;
Matches 83; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

212 aactccaaatctaacagactctgactccgcttcttgatgaacttgcagccaatgctca 271
|||||
83 AACTCCAAATCTAACAGACTCTGACTCCGCTTGTGATGAACCTGGCCAGCCAATGCTCA 24

QY 272 agtatcaagggtttacacactgga 294
Db 23 AGTATCAAGGTTTACACACTGGA 1

RESULT 7
AAZ59035/C
ID AAZ59035 standard; DNA; 83 BP.
XX AC AAZ59035;
XX XX
XX DT 11-APR-2000 (first entry)
XX DE Oligonucleotide A-5 for generating codon-optimised ferric reductase.
XX KW Crop plant; polyadenylation signal; substitution; ferric reductase; iron;
XX KW yeast; heterologous gene; ss.
XX OS Saccharomyces cerevisiae.
XX OS Synthetic.
XX PN WO9948356-A1.

XX OS Saccharomyces cerevisiae.
XX OS Synthetic.
XX PN WO9948356-A1.
XX PD
XX PF
XX PI
XX PR
XX PP
XX PT
XX PS
XX CC The invention relates to a method of transforming a higher plant with
XX CC a foreign gene in which the polyadenylation signal region contained in
XX CC the nucleotide sequence of the foreign gene has been modified by
XX CC nucleotide substitution to give a sequence not having this function, in
XX CC such a way that the function of the foreign gene expression product is
XX CC not substantially impaired. The foreign gene used is especially an iron
XX CC (Fe) 3+ (ferric) reductase gene in which putative polyadenylation
XX CC signals contained within the coding sequence are substituted with
XX CC alternative codons for expression in higher plants. Oligonucleotides
XX CC AAZ59031-259060 were used as overlapping oligonucleotide primers to
XX CC generate a plant codon-optimised yeast ferric reductase (FREL gene;
XX CC AAZ59028) in which putative polyadenylation signal and poly(A) addition
XX CC sites are removed by nucleotide substitution. The method allows the
XX CC production of transformant plants with improved expression of a desired
XX CC foreign gene, especially to improve iron absorption efficiency in crop
XX CC plants.
XX SQ Sequence 83 BP; 18 A; 15 C; 20 G; 30 T; 0 other;

Query Match 4.0%; Score 83; DB 20; Length 83;
Best Local Similarity 100.0%; Pred. No. 5.2e-16;
Matches 83; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

282 ttatcacactggaggacatgaagaacatctaccttaataatgaagtaactacattcgctc 341
|||||
83 TTTACACTGGAGGACATGAAGAACATCTACCTTAATGAAGTAACTTACCTTCGCTC 24

QY 342 ctgagaaatccgataaagagaca 364
Db 23 CTGAGAAATCCGATAAAGAGACA 1

RESULT 8
AAZ59036/C
ID AAZ59036 standard; DNA; 83 BP.
XX AC AAZ59036;
XX XX
XX DT 11-APR-2000 (first entry)
XX DE Oligonucleotide A-6 for generating codon-optimised ferric reductase.
XX KW Crop plant; polyadenylation signal; substitution; ferric reductase; iron;
XX KW yeast; heterologous gene; ss.
XX OS Saccharomyces cerevisiae.
XX OS Synthetic.
XX PN WO9948356-A1.
XX PD
XX PF
XX PI
XX PR
XX PP
XX PT
XX PS
XX CC The invention relates to a method of transforming a higher plant with
XX CC a foreign gene in which the polyadenylation signal region contained in
XX CC the nucleotide sequence of the foreign gene has been modified by
XX CC nucleotide substitution to give a sequence not having this function, in
XX CC such a way that the function of the foreign gene expression product is
XX CC not substantially impaired. The foreign gene used is especially an iron
XX CC (Fe) 3+ (ferric) reductase gene in which putative polyadenylation
XX CC signals contained within the coding sequence are substituted with
XX CC alternative codons for expression in higher plants. Oligonucleotides
XX CC AAZ59031-259060 were used as overlapping oligonucleotide primers to
XX CC generate a plant codon-optimised yeast ferric reductase (FREL gene;
XX CC AAZ59028) in which putative polyadenylation signal and poly(A) addition
XX CC sites are removed by nucleotide substitution. The method allows the
XX CC production of transformant plants with improved expression of a desired
XX CC foreign gene, especially to improve iron absorption efficiency in crop
XX CC plants.
XX SQ Sequence 83 BP; 18 A; 15 C; 20 G; 30 T; 0 other;

Transformation of higher plant with foreign gene having modified
poly(A) site -
Disclosure; Fig 6; 8lpp; Japanese.
(NISC-) JAPAN SCI & TECHNOLOGY CORP.
Mori S, Nakanishi H, Oki H, Yamaguchi H;
WPI; 1999-571924/48.

```


XX Disclosure; Fig 6; 8lpp; Japanese.
 XX The invention relates to a method of transforming a higher plant with
 CC a foreign gene in which the polyadenylation signal region contained in
 CC the nucleotide sequence of the foreign gene has been modified by
 CC nucleotide substitution to give a sequence not having this function, in
 CC such a way that the function of the foreign gene expression product is
 CC not substantially impaired. The foreign gene used is especially an iron
 CC (Fe) 3+ (ferric) reductase gene in which putative polyadenylation
 CC signals contained within the coding sequence are substituted with
 CC alternative codons for expression in higher plants. Oligonucleotides
 CC AAZ59031-259060 were used as overlapping oligonucleotide primers to
 CC generate a plant codon-optimised yeast ferric reductase (FRE1 gene;
 CC AAZ59028) in which putative polyadenylation signal and poly(A) addition
 CC sites are removed by nucleotide substitution. The method allows the
 CC production of transformant plants with improved expression of a desired
 CC foreign gene, especially to improve iron absorption efficiency in crop
 CC plants.
 XX Sequence 83 BP; 18 A; 21 C; 18 G; 26 T; 0 other;
 SQ

Query Match 4.0%; Score 83; DB 20; Length 83;
 Best Local Similarity 100.0%; Pred. No. 5.2e-16;
 Matches 83; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 911 gcatacgtctgttcattgtagcgtgcgtccattcaatcgttatgacgcgttcaggagtt 970
 Db 1 gcatacgtctgttcattgtagcgtgcgtccattcaatcgttatgacgcgttcaggagtt 60
 QY 971 aaacgagagattccagtcctc 993
 Db 61 aaacgagagattccagtcctc 83

RESULT 11
 AAZ59045
 ID AAZ59045 standard; DNA; 83 BP.
 XX
 AC AAZ59045;
 XX
 DT 11-APR-2000 (first entry)
 DE Oligonucleotide C-4 for generating codon-optimised ferric reductase.
 KW Crop plant; polyadenylation signal; substitution; ferric reductase; iron;
 yeast; heterologous gene; ss.
 OS Saccharomyces cerevisiae.
 XX Synthetic.
 PN WO9948356-A1.
 XX
 PD 30-SEP-1999.
 XX
 PF 24-MAR-1999; 99WO-JP01481.
 XX
 PR 24-MAR-1998; 98JP-0096637.
 XX
 PA (NISC-) JAPAN SCI & TECHNOLOGY CORP.
 XX
 PI Mori S, Nakanishi H, Oki H, Yamaguchi H;
 XX WPI; 1999-571924/48.
 XX
 CC Transformation of higher plant with foreign gene having modified
 PT poly(A) site -
 XX
 PS Disclosure; Fig 6; 8lpp; Japanese.
 XX The invention relates to a method of transforming a higher plant with
 CC a foreign gene in which the polyadenylation signal region contained in

CC the nucleotide sequence of the foreign gene has been modified by
 CC nucleotide substitution to give a sequence not having this function, in
 CC such a way that the function of the foreign gene expression product is
 CC not substantially impaired. The foreign gene used is especially an iron
 CC (Fe) 3+ (ferric) reductase gene in which putative polyadenylation
 CC signals contained within the coding sequence are substituted with
 CC alternative codons for expression in higher plants. Oligonucleotides
 CC AAZ59031-259060 were used as overlapping oligonucleotide primers to
 CC generate a plant codon-optimised yeast ferric reductase (FRE1 gene;
 CC AAZ59028) in which putative polyadenylation signal and poly(A) addition
 CC sites are removed by nucleotide substitution. The method allows the
 CC production of transformant plants with improved expression of a desired
 CC foreign gene, especially to improve iron absorption efficiency in crop
 CC plants.
 XX Sequence 83 BP; 22 A; 19 C; 13 G; 29 T; 0 other;
 SQ

Query Match 4.0%; Score 83; DB 20; Length 83;
 Best Local Similarity 100.0%; Pred. No. 5.2e-16;
 Matches 83; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 981 tattccagtcctctgttaaggaaattctacttcagatgggggaatagtagccacaattctta 1040
 Db 1 tattccagtcctctgttaaggaaattctacttcagatgggggaatagtagccacaattctta 60
 QY 1041 tgtccatcatcatttccagtc 1063
 Db 61 tgtccatcatcatttccagtc 83

RESULT 12
 AAZ59046/C
 ID AAZ59046 standard; DNA; 83 BP.
 XX
 AC AAZ59046;
 XX
 DT 11-APR-2000 (first entry)
 DE Oligonucleotide C-4 for generating codon-optimised ferric reductase.
 KW Crop plant; polyadenylation signal; substitution; ferric reductase; iron;
 yeast; heterologous gene; ss.
 OS Saccharomyces cerevisiae.
 XX Synthetic.
 PN WO9948356-A1.
 XX
 PD 30-SEP-1999.
 XX
 PF 24-MAR-1999; 99WO-JP01481.
 XX
 PR 24-MAR-1998; 98JP-0096637.
 XX
 PA (NISC-) JAPAN SCI & TECHNOLOGY CORP.
 XX
 PI Mori S, Nakanishi H, Oki H, Yamaguchi H;
 XX WPI; 1999-571924/48.
 XX
 CC Transformation of higher plant with foreign gene having modified
 PT poly(A) site -
 XX
 PS Disclosure; Fig 6; 8lpp; Japanese.
 XX The invention relates to a method of transforming a higher plant with
 CC a foreign gene in which the polyadenylation signal region contained in
 CC the nucleotide sequence of the foreign gene has been modified by
 CC nucleotide substitution to give a sequence not having this function, in
 CC such a way that the function of the foreign gene expression product is
 CC not substantially impaired. The foreign gene used is especially an iron
 CC (Fe) 3+ (ferric) reductase gene in which putative polyadenylation

SQ Sequence 83 BP; 23 A; 20 C; 23 G; 17 T; 0 other;

Query Match 4.0%; Score 83; DB 20; Length 83;
Best Local Similarity 100.0%; Pred. No. 5.2e-16;
Matches 83; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1193 gctggcattctctgttcgacaggttctgcgaattgtacgtatcatcatgaacggaggt 1252
|||||
Db 83 GCTGGCATTCTCTGCTTCGACAGGTTCTGCCGATTGTACGTAATCATCATGAACGGAGGT 24

Qy 1253 cttgaagccgcacctttgtcgac 1275
|||||
Db 23 CTTAAGACCGCCACTTTGTGCGAC 1

RESULT 15
AAZ59049
ID AAZ59049 standard; DNA; 82 BP.

AAZ59049;

DT 11-APR-2000 (first entry)

XX Oligonucleotide D-1 for generating codon-optimised ferric reductase.

DE Crop plant; polyadenylation signal; substitution; ferric reductase; iron;
KW yeast; heterologous gene; ss.

XX Saccharomyces cerevisiae.

OS Synthetic.

XX WO9948356-A1.

XX 30-SEP-1999.

XX 24-MAR-1999; 99WO-JP01481.

XX 24-MAR-1998; 98JP-0096637.

XX (NISC-) JAPAN SCI & TECHNOLOGY CORP.

XX Mori S, Nakanishi H, Oki H, Yamaguchi H;

XX WPI; 1999-571924/48.

PT Transformation of higher plant with foreign gene having modified
poly(A) site -

Disclosure; Fig 6; 81pp; Japanese.

XX The invention relates to a method of transforming a higher plant with
CC a foreign gene in which the polyadenylation signal region contained in
CC the nucleotide sequence of the foreign gene has been modified by
CC nucleotide substitution to give a sequence not having this function, in
CC such a way that the function of the foreign gene expression product is
CC not substantially impaired. The foreign gene used is especially an iron
CC (Fe) 3+ (ferric) reductase gene in which putative polyadenylation
CC signals contained within the coding sequence are substituted with
CC alternative codons for expression in higher plants. Oligonucleotides
CC AAZ59031-259060 were used as overlapping oligonucleotide primers to
CC generate a plant codon-optimised yeast ferric reductase (FRE1 gene;
CC AAZ59028) in which putative polyadenylation signal and poly(A) addition
CC sites are removed by nucleotide substitution. The method allows the
CC production of transformant plants with improved expression of a desired
CC foreign gene, especially to improve iron absorption efficiency in crop
CC plants.

SQ Sequence 82 BP; 24 A; 18 C; 17 G; 23 T; 0 other;

Query Match 3.9%; Score 82; DB 20; Length 82;
Best Local Similarity 100.0%; Pred. No. 1.1e-15;

Matches 82; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1270 gtcgaccacagatgattcttaacgtttatcaagatctctgtcaagaagccttaagttcttcaa 1329
|||||
Db 1 gtcgaccacagatgattcttaacgtttatcaagatctctgtcaagaagccttaagttcttcaa 60

Qy 1330 gtatcaagtgggagcatttgcc 1351
|||||
Db 61 gtatcaagtgggagcatttgcc 82

Search completed: May 14, 2002, 07:27:53
Job time: 5730 sec



GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 14, 2002, 04:19:53 ; Search time 2978.83 seconds
(without alignments)
11585.790 Million cell updates/sec

Title: US-09-646-825-1
Perfect score: 2092
Sequence: 1 gaattctctagactccacca.....cttggtaagagctcaagctt 2092

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0
Searched: 1472140 seqs, 8248589755 residues
Total number of hits satisfying chosen parameters: 2944280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl :
1: gb_ba :
2: gb_htg :
3: gb_in :
4: gb_om :
5: gb_ov :
6: gb_pat :
7: gb_ph :
8: gb_pl :
9: gb_pr :
10: gb_ro :
11: gb_sts :
12: gb_sy :
13: gb_un :
14: gb_vi :
15: em_ba :
16: em_fun :
17: em_hum :
18: em_in :
19: em_om :
20: em_or :
21: em_ov :
22: em_pat :
23: em_ph :
24: em_pl :
25: em_ro :
26: em_sts :
27: em_sy :
28: em_un :
29: em_vi :
30: em_htgo_hum :
31: em_htgo_inv :
32: em_htgo_rod :
33: em_htgo_hum :
34: em_htg_inv :
35: em_htg_rod :
36: em_htg_other :

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	2092	100.0	2092	6	E31006	E31006 Method for M86908 S. cerevisia
2	1244	59.5	3264	8	YSCFRE1A	U14913 Saccharomyc
3	1244	59.5	38868	8	YSCHE167	A86082 Sequence 74
4	147.6	7.1	694	6	A86082	ARI55575 Sequence
5	147.6	7.1	694	6	ARI55575	E66100 Genome DNA
6	147.6	7.1	694	6	E66100	D83181 Candida alb
7	99.4	4.8	3084	8	D83181	AJ387722 Candida a
8	97.8	4.7	3615	8	CAL387722	ARI5974 Sequence 63
9	90.2	4.3	711	6	ARI5974	E65992 Genome DNA
10	90.2	4.3	711	6	ARI55467	E31009 Method for
11	90.2	4.3	711	6	E65992	E31010 Method for
12	83	4.0	83	6	E31009	E31011 Method for
13	83	4.0	83	6	E31010	E31012 Method for
14	83	4.0	83	6	E31011	E31013 Method for
15	83	4.0	83	6	E31012	E31014 Method for
16	83	4.0	83	6	E31013	E31021 Method for
17	83	4.0	83	6	E31014	E31022 Method for
18	83	4.0	83	6	E31021	E31023 Method for
19	83	4.0	83	6	E31022	E31024 Method for
20	83	4.0	83	6	E31023	E31025 Method for
21	83	4.0	83	6	E31024	E31026 Method for
22	83	4.0	83	6	E31025	E31027 Method for
23	83	4.0	83	6	E31026	E31028 Method for
24	82	3.9	82	6	E31027	E31029 Method for
25	82	3.9	82	6	E31028	E31030 Method for
26	82	3.9	82	6	E31029	E31031 Method for
27	82	3.9	82	6	E31030	E31032 Method for
28	82	3.9	82	6	E31031	E31033 Method for
29	82	3.9	82	6	E31032	E31034 Method for
30	80	3.8	80	6	E31015	E31035 Method for
31	80	3.8	80	6	E31016	E31036 Method for
32	80	3.8	80	6	E31017	E31037 Method for
33	80	3.8	80	6	E31018	E31038 Method for
34	80	3.8	80	6	E31019	E31039 Method for
35	80	3.8	80	6	E31020	E31040 Method for
36	77	3.7	77	6	E31033	E31041 Method for
37	77	3.7	77	6	E31034	E31042 Method for
38	77	3.7	77	6	E31035	E31043 Method for
39	77	3.7	77	6	E31036	E31044 Method for
40	77	3.7	77	6	E31037	E31045 Method for
41	77	3.7	77	6	E31038	E31046 Method for
42	63	3.0	3987	8	SCYOR384W	E31047 Method for
43	58.8	2.8	6955	8	SCYRL220C	E31048 Method for
44	58.8	2.8	13213	8	SCSORE	E31049 Method for
45	53	2.5	974	11	CNS061Z7	E31050 Method for

ALIGNMENTS

RESULT	1
E31006	E31006
LOCUS	E31006
DEFINITION	Method for transforming plant, the thus transformed plant and gene thereof.
ACCESSION	E31006
VERSION	E31006.1 GI:13021492
KEYWORDS	JP 1999266876-A/1.
SOURCE	unidentified.
ORGANISM	unclassified.
REFERENCE	1 (bases 1 to 2092)
AUTHORS	Satoshi, M.H.O.O. and Nakanishi, H.Y.
TITLE	Method for transforming plant, the thus transformed plant and gene thereof
JOURNAL	Patent: JP 1999266876-A 1 05-OCT-1999;
COMMENT	SCIENCE & TECH AGENCY
	OS Unidentified
	PN JP 1999266876-A/1
	PD 05-OCT-1999

```

PF 24-MAR-1998 JP 1998096637
PR SATOSHI MORI,HIROYUKI OKI,NOBUHITO NAKANISHI, PI HIROTAKE
YAMAGUCHI
PC C12N15/09,A01H5/00,C12N5/10,C12P19/34//C12N15/09,C12R1:645),
PC (C12N5/10,C12R1:91),C12N15/00,C12N5/00,(C12N15/00,C12R1:645),
PC (C12N5/00,C12R1:91)
CC Strandedness: Both;
CC Topology: Linear;
FH key Location/Qualifiers
FT source 1..2092 /organism='unidentified'.
FEATURES
    source
        Location/Qualifiers
            1..2092 /organism='unidentified'
            /db_xref="taxon:32644"
BASE COUNT 587 a 510 c 418 g 577 t
ORIGIN
Query Match 100.0%; Score 2092; DB 6; Length 2092;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2092; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 gaattcttagactccaccatggtagaacagagtccttttctgctcttctcattcttt 60
Db 1 GAATCTCTAGACTCCACCATGGTTAGAACAGAGTCCTTTTCTGCCTCTTCATCTCTT 60
Qy 61 ctctgctacagtcacatcgagcgctacactcactcactcactcactcactcactcactc 120
Db 61 CTTCGGCTACAGTCCAAATCGAGCGCTACACTCATCTCCACTTCATGCCATTTCTCAGGTCG 120
Qy 121 actgtaccagttcgatgctcaagcaagtcacaaagtccttctgactgcaagaacacaaattg 180
Db 121 ACTGTACAGTTCGGATGCTCAAGCAAGTCAAAAGTCTTGCTACTGCAAGAACATCAATTG 180
Qy 181 gctcggaagcgctcactgcactgcgtttagaagactccaaatcctaaagactctggactc 240
Db 181 GCTCGGAAGCGTCACTGCATGCGCTTATGAGAACTCCAAATCTAAAGACTCTGGACTC 240
Qy 241 cgttttgatgaacttcgagcgaatcctcaagtatcaaggtttacacactggaggacat 300
Db 241 CGCTTTGATGAATTTGCCAGCCCAATGCTCAAGTATCAAGGTTTACACACTGGAGGACAT 300
Qy 301 gaagaacatctacatttaagtcaagtaactacctcgctcctcgctgagaaatccgataagaa 360
Db 301 GAAGAACATCTACCTTAATGCAAGTAATACTTCCGCTCTCGAGAAATCCGATAGAA 360
Qy 361 gacagttgtttcacacacgcttgatgagcaaatgagacgacctatcactactactatgagga 420
Db 361 GACAGTTGTTTACACACCGTTGATGGCAATGAGACGCGCTTATCATTACTATGAGGA 420
Qy 421 aaactatgggatacacttgaaattgattgactcactcaatggtgagcatggggcctcgcttt 480
Db 421 AAATATGGGATCCACTTGAAATTTGATGCGATCTCAATGGTGGCATGGGCGCTCGTCTT 480
Qy 481 ctctgggtcgagtccttaccgcgcgaactatcttgaacattcctcaaacgctattcgg 540
Db 481 CTCTGGGTGCGACTCTTACCAGCGCAACTATCTTGNACATTTCTCAACCGCTATTGCG 540
Qy 541 caagaacattatggcaaatcttctgtaagaagtccttctatctaccacgagcgtttacaaga 600
Db 541 CAAGAACATTATGGCAAAATCTGTTAAGMAGTCTTATCTATACCGCAAGCGCTTACAAGA 600
Qy 601 ctacaagagagaactttctatctttggaaacggtttgcccattcaacttcaaacctcgagg 660
Db 601 CTACAAGAGAGAATTTCTATCTTTGGAAACGTTTGCCATTCAACTTTACAACCTCGAGG 660
Qy 661 caaaggactcgtagttcttcttctgctacactgactattctcactctcttcttcggaca 720
Db 661 CAAAGGACTCGTAGTCTTATCTTTGTTCATCTGACTATTCTCTCCTCTCTTCGGACA 720
Qy 721 taacatcaagttgccaactccttacgtagacccatagatggagaagatacaatggcattcgt 780

```

```

Db 721 TAACATCAAGTTGCCACATCCTTTACGATAGACCTTAGATGAGAGAGATCAATGGCATTTCGT 780
Qy 781 ctcaagccgctgactgactggtgcaactcgctcttctcccggtggtgctacettctcggtat 840
Db 781 CTACAGCCGCTGCTGACTGATGATGGAATCGCTCTTTTCCCGGTGCTGACCTTTTCGGTAT 840
Qy 841 ccggaacaaccccttcatcccaatccacggaattgagcttttagtacttctcaacttttacc 900
Db 841 CCGGAACAACCCCTTCATCCCAATCACGGATTGAGCTTTAGTAGTACTTTCAACTTTTACCA 900
Qy 901 caaatggtcagcatacctcgtcttcattgtagcgttcgtccattcaactcgttatgacacgc 960
Db 901 CAAATGCTAGCATACCTGCTTCTTATGTCATGTTAGCGCTGCTCCATTCATTCGTTATGACCG 960
Qy 961 ttcaggagttaaacgagagagattccagtcctctgtaaggaattcttacttcagatgggg 1020
Db 961 TTCAGGAGTTAAACGAGGAGTATTCAGTCTCTTGTAGGAAATTCATCTCAGATGGGG 1020
Qy 1021 aatagtagccacaattcttatgtccatcacttctccagtcggagaggtcttcaggaa 1080
Db 1021 AATAGTAGCCACATCTTATGTCATCATCATCTTTTCCAGTCCGAGAAGGCTTTTCAGGAA 1080
Qy 1081 ccgaggttatgaattcttcttacttatccaaagccatgaacatcatgtttatcatagc 1140
Db 1081 CCGAGGTTATGAATCTTCTTACTTATTCACAAAGCCATGAACATCATGTTTATCATAGC 1140
Qy 1141 tatgtattacattgcccacacactagatggctgggctggttcctcctggtcgtggcat 1200
Db 1141 TATGTATTACATTGCCCACACACTAGGATGGATGGGCTGGATCTGGTCCATGGTGGCAT 1200
Qy 1201 cctctgttcgacaggttctgcggaattgtacgtatcatcatcgaacgagagctcctaagac 1260
Db 1201 CCTCTGTTGCACAGGTTCTGCCGAATGTACGTATCATCATGAACGAGGCTCTTAAGAC 1260
Qy 1261 cgcactttgtcgaccacagatgattctaaagttatcaagatctctgtcgaagaagcctaa 1320
Db 1261 CGCCACTTTGTCGACCACAGATGATTTACGTTATCAAGATCTCTGTGTCAGAGAGCCCTAA 1320
Qy 1321 gttctcaagtatcaagtgaggagcatttgcctatgatttactttcttccacaaatcagc 1380
Db 1321 GTTCTTCAAGTATCAAGTGGGAGCATTTGCCCTATATGTACTTTCTTCCACCAAAATCAGC 1380
Qy 1381 ctggtctacagtttcaatctcctcctcactcactcactcactcactcactcactcactc 1440
Db 1381 CTGGTTCTACAGTTTCAATCTCATCTCCCTTCCAGTCTCTATCAGAAAGGACACAGATCC 1440
Qy 1441 taacaacccagatcaactaactatgtaogtcaaaagctaaacagggcattagagagact 1500
Db 1441 TAAACAACCCAGATCAACTTATGTAGCTCAAGCTAACAGGGCATTCAGAGAGTACT 1500
Qy 1501 tcttagcaagttcttaagcgtcccaaacctacccgttgcattgcaagatttctcttagaggg 1560
Db 1501 TCTTAGCAAAAGTTCTAAGCGCTCCAAACCATACCGTTGATTGCAAGATTTTCTTAGAGGG 1560
Qy 1561 accataggcgtaactgtccctcacttgcacaaacttaagagaatcttagtagagtagc 1620
Db 1561 ACCATATGGCGTAACTGTCCCTCATATTGCCAAACTTAAGAGAAATCTAGTAGGAGTAGC 1620
Qy 1621 tgcgggctcggtggcgagcacttaccgccatttctgtagaagtccttagattgacctag 1680
Db 1621 TGCGGGCTCGCGGTGGCAGCCATCTACCCCATTTCTAGTAATGCCCTTAGATTGCCCTAG 1680
Qy 1681 cactgatacaactgcagacaagttctactggttcacgcaccccttagctacacttaagtg 1740
Db 1681 CACTGATCAACTGAGCAGCAAGTTCTACTGGATCGTCAACGACCTTAGTACCTTAAGTG 1740
Qy 1741 gttcgaacacagctacaatggcttaaggagaatctctgtaagtcctctgtaactacac 1800
Db 1741 GTTCGAAACAGAGCTACAATGGCTTAAGGAGAAATCTTGTGAAGTCTCTGTCACTATCAGAC 1800
Qy 1801 tgggtcactgagggagatatacaaaactcagatgagtcceactaagggttttcgagtagaca 1860

```

Db 1801 TGGGTCATCAGTGGAGGATACAAACTCAGATGAGTCCACTAGGCGTTTCGATGACAAGGA 1860

QY 1861 aqaatctgaatcacgtagaatgcttaacaagagccagacacctcaagaagcagtagtgaq 1920

Db 1861 AGAATCTGAATCACCAGTAGAATGCTTTAACAAAGAGCCAGACCTCAAGAGCTAGTAGAG 1920

QY 1921 atcagagatcaaatgtcagaactcgagaacacaaacacacactcttctactcatcgcgacc 1980

Db 1921 ATCAGAGATCAAAATGTCAGAACTCGAGAACCAACAACATCACTTCTACTCATCGGACC 1980

QY 1981 aagcactttcaatgacgacttaggaatgcagttgataaaggatcagatttagtctgaa 2040

Db 1981 ACGGACTTTCAATGACGACTTTAGGAATGCGAGTTGTACAAGGATATCGAATCTAGTCTGAA 2040

QY 2041 gatagatgcgaactagaggagaggttttacttggttaagagctcaagctt 2092

Db 2041 GATAGATGCGAACTAGAGGAGAGAGTTTACTTGGTAAGAGCTCAAGCTT 2092

RESULT 2

YSCFREIA 3264 bp DNA PLN 27-APR-1993

DEFINITION S. cerevisiae ferric reductase gene, complete cds.

ACCESSION M86908

VERSION M86908.1 GI:171520

KEYWORDS ferric reductase.

SOURCE Saccharomyces cerevisiae DNA.

ORGANISM Saccharomyces cerevisiae

REFERENCE 1 (bases 1 to 3264)

AUTHORS Dancis,A., Roman,D.G., Anderson,G.J., Hinnebusch,A.G. and Klausner,R.D.

TITLE Ferric reductase of saccharomyces cerevisiae: Molecular characterization, role in iron uptake, and transcriptional control by iron

JOURNAL Proc. Natl. Acad. Sci. U.S.A. 89, 3869-3873 (1992)

MEDLINE 92237270

FEATURES

source 1..3264

Location/Qualifiers

/organism="Saccharomyces cerevisiae"

/db_xref="taxon:4932"

1..977

/gene="FRE1"

/note="confers iron regulated activity to heterologous coding region"

1..3264

/gene="FRE1"

929..3264

/note="5' end mapped by primer extension, 3' end estimated from mRNA size"

978..3038

/gene="FRE1"

/note="essential for ferric reductase activity"

/codon_start=1

/product="ferric reductase"

/protein_id="AAA34608.1"

/db_xref="GI:171521"

translation="MTRVRLVFLCLFISFFATVQSSATLISQIAALYQGCSSKS KSCYCNINWLSGVTVACAYENSKNTLDSALMKLASQSSIKVYTLDDMKNIYLNAS NYLRAPESDKRTVYVQPLMANETAYHYVEENYGIHLNLMRSOWCAWGLVFWAVL TAATLNLKRVGKNIANSVKSLIYPSYKDYNTERTFVLMKRLPENFTRGKGLV VLIFVILILSLSFHGNLKLPHVDRPRWRMSAFVSRADLMALPFPVYLGIKRN NPFDITGLSESTFNFYHKSAYVCFMLAVVHSIVNTASGVRGVFSVLVRKFFYRFG IVATLTIISIFQSKVFNRRGYEFLIHKAMNIMFIAMYYHCTGLWGMWISMA GTLCDFRCRIVRIIMNGLKLATLSTDDSNVIKISKPKFFKYQVGFAYMYFLS PKSAWYFQSPHPFTVLSERHDPNNPDLTMVKANGITRVLKSLVLSAPNHPVDC KIFLPGYGVTPHTAKLRNLVGAAGLVAAIYPHEVECLRLPSTDLQHKFVIW NDLSHLKWFENELWLKEKSEVSIYTGSSVEDNSDESTKGFDDKRESEITVECLN KRPDLKELVLRSEIWLKLENNNITFYSCGPATFNDFRNAVVGIDLSLKIDVELEE SFTW"

BASE COUNT 971 a 591 c 630 g 1072 t

ORIGIN

Query Match 59.5%; Score 1244; DB 8; Length 3264;

Best Local Similarity 75.0%; Pred. No. 0;

Matches 1556; Conservative 0; Mismatches 520; Indels 0; Gaps 0;

QY 16 caccatggttagaaccagagtcctttctgcctctctcattctctctctcgtcacagtcaca 75

Db 974 CAATATGTTAGAACCGGTGATATTCTGCTATTATATCTTTTGTGCTACGGTTCA 1033

QY 76 atcagagctacactcatctccacttcattgcattctcaggtcagctgacagtacagttcgg 135

Db 1034 ATCGAGTGTACACTTATTAGCACTTCATGTATTTCCCAAGCTGCGCTATACCAATTTGG 1093

QY 136 atgctcaagcaagtcgaagcttctgactgcaagaacatcaattgctcgggaagcgtcac 195

Db 1094 ATGTTCTTAGTAAATCTAAAGTTGCTACTGTAAACAACTCAATTTGGTTCAGTGAC 1153

QY 196 tqcatgcttatgagaactccaaatctaaagaactctggagactcgcgtcttgatgaaact 255

Db 1154 AGCATGTGCTATGAGAAATTCCAATCTAACAAACACTAGACAGGCGCTTAATGAAGTT 1213

QY 256 tgccagcgaatgctcaagtatcaaggttttacacactggaggacatgaagaacatctacat 315

Db 1214 AGCATCCCAATGTTCAAGCATCAAAGTTTATCTTTAGAGGACATGAAGAATATTATTT 1273

QY 316 taatgcaagtaactaccctcgcgtcctcgtgagaaatccgataagaagacagtggtttcaca 375

Db 1274 AAATGCGTCAAAATTTATTTGAGAGCACCTGAGAAAAAGTATAAAAAACCGTGTAGTCA 1333

QY 376 accgttgatgcaaatgagagcgccctatcactactactatgagaaactatggatcca 435

Db 1334 ACCGTCATGGGACGAGACAGCGATATCTATTATTATGAGGAAAAATTTATGTTATCCA 1393

QY 436 cttgaatttgatcgatctcaatggtgcgcgtggggcctcgtctctctctctgggtgcagt 495

Db 1394 TCTTAACCTAATGCGCTCCTCAATGTTGGTGGCTTCGCTCTTCTTCTGGTGGCTGT 1453

QY 496 ccttaccgcgcgaactatctgaacattctcaaacgctattcgcgcaagaacattatgac 555

Db 1454 GCTTATGCGAGCCACTATCTTGAACATTTCTGAAAAAGGGTGTGTTGTAAGAACATCATGC 1513

QY 556 aaatctgttaagaagctctctatctaccacagcgcttacaagactacaacagagagaaac 615

Db 1514 AAATCCGTCAAAAATCACTTATTATCTCTGCTTTTACAAGATTTATTAATGAAGCAAC 1573

QY 616 tttctatctttggaaacgctttgcccattccaactccaactcagcagcaggaagcagtagt 675

Db 1574 TTTTATTTATGGAAGCGCTCTACCATTTAATTTTACAACCTGAGGCAAGGGTCTCGCTGT 1633

QY 676 tottatctttgctcattctgactattctcactctctctctctctcggacataacatacagttgc 735

Db 1634 ATTAATTTTGTATTATTGACTATATATTCTCTCAAGTTTGGTCAATAATTAATAACTCC 1693

QY 736 acatcctcatgataccttagatgagaagatcaatggcattcgtctcagcgcgtagtga 795

Db 1694 ACACCATATGATAGCCCGAGATGGAGAAAGATAGGCGCTTTGTGAGCTGTAGACAGCA 1753

QY 796 cttgatggcaatcgctctcttcccgcgtggtgacattcttcgttatccggaacaacccctt 855

Db 1754 CTTGATGGCCATTGGCACTTTTCCAGTAGCTCTATCTATTTCGGAATAAGAAATAATCCCTT 1813

QY 856 catcccaatcaccgagattgagcttagttagttagttagttagttagttagttagttagttag 915

Db 1814 CATCCCTATAACAGGGCTTTTCCCTTTTCTACATTTAATTAATTAATTAATTAATTAATTA 1873

QY 916 cgtcgtctcatgttagcgtcgtccattcaatcgtagcagcgttcagaggttaaacg 975

Db 1874 CGTTGTTTTCATGTTGGCGCTTGTGACACTCAATTTGATGACCGCTCGGGAGTGAAG 1933

QY 976 aggagTattccagctctctgttaaggaaattctacttcagatggggaatagtagcccaat 1035

Db	30307	ATGTTCTAGTAAATCTAAAGCTTGCTACTGTGTAAGAACATCAATTTGGCTGGGTTTCAGTGCAC	30366
Qy	196	tgcctgcgcttatgagaactccaaactcaacgaactcttgactccgcttttgatgaagaaact	255
Db	30367	AGCATGTGCCCTATGAGAATTCCAATCTAACAAACACTAGACAGCGCCTTAATGAAGTT	30426
Qy	256	tgcagcccaatgcctcaagtatcaaggtttacacactggaggacatgaagaacatctacct	315
Db	30427	AGCATCCCAATGTTCAAGCATCAAAGTTTATCTTTAGAGGACATGAAGAATATTTATTT	30486
Qy	316	taatgcaagttaactacctctcgcgctcctcgagaataacgataagaagacagttgtttcaaca	375
Db	30487	AAATGCGTCAAAATATTTGAGAGCACCCTGAGAAAGTGATAAAAAACCCGGTGTAGTCA	30546
Qy	376	accgttgatggcaaatgagacgcgctactactactatgagaaaactatggatcca	435
Qy	30547	ACCGCTCATGGCAACGAGACAGCGTATCATTTATTTATGAGGANAATTTATGTTATCCA	30606
Qy	436	cttgaattgtagcgaatcctaatggtagcgaatggggcctcgtctcttcttggtgcgagt	495
Qy	30607	TCTTAAACCTAATGGCTCTCAATGGTGGCTTGGGCTCTGCTCTTCTTCTGGTGGCTGT	30666
Qy	496	ccttacgcgcgaactatcttgaacattctcaaacgcgatatcggcaagaacattatggc	555
Db	30667	GCTTACTCGACGCCACTATCTTTGAACATTTCTGAAAGGGTGTTGGTAAGAACAATCATGGC	30726
Qy	556	aaattctgttaagaagtccttactctaccgaacgctttacaagaactacaacgagagaac	615
Db	30727	AAACTCCGTCAAAAATCATCTATTTATCCCTCTGTTTACAAAGATTTATAATGAACGAAAC	30786
Qy	616	ttctctatcttggaaaacgcttggcattccaactctacaactcggagcgaagactcgtagt	675
Db	30787	TTTTTATTTATGGAAGCGCTACCATTTAAATTTAGAACTCGAGCAAGGCTCTCGTCGT	30846
Qy	676	tcttatcttbtgcattctgcatactctcactctctcttcggacataaacatcgaagtgc	735
Db	30847	ATTAATTTTTTGTATTTTGGACTATATATATCTCTCAGTTTGGTCATATAATTAACACTCC	30906
Qy	736	acatccttacatagacctagatgagaagacatcaatggcatctgctcacgcctgctga	795
Db	30907	ACACCATATGATAGGCCAGATGAGGAAGATATGGCCCTTGTGAGTCTGTAGACGAGA	30966
Qy	796	cttgatggcaatcgcctctttcccgcgtggtgcactcttcggtatcccggaaacacccctt	855
Db	30967	CTTGATGCCCATTGCACATTTTCCACGATGCTATCTATTTCGGAATAAGAAATAATCCCTT	31026
Qy	856	catcccaatcacccgattgagccttagtactttcaactttcaccaaatggtcagcata	915
Qy	31027	CATCCCTATAACAGGGCTTTCCTTTCTCFACATTTTAAATTTCTATCATAAATGGTCTGCCTA	31086
Qy	916	cgtctgcttcagttagccgcgtccactcgaatcgttatgcaccttcagagattaaacg	975
Db	31087	CGTTTGTGTTTCATGTTGGCCGTTGTACACTCAATTTGTCATGACCCGCTCGGGAGGTGAAAG	31146
Qy	976	aggagattccagtccttctgtaaggaattctacttcagatggggaaatagtagccacaat	1035
Db	31147	AGGTGTGTTTCAAGTCTGGTTAGGAATTTTACTTTAGTGGGGTATAGTGGCAACGAT	31206
Qy	1036	tcttatgtccatcatcatcttccagtcgcgagaaggtcttcaggaaacccaggtttatgaaat	1095
Db	31207	ATTAATGTCTATTTATTTTCCAAAGTGAAAAAGTATTTAGAAATAGAGGGTATGAGAT	31266
Qy	1096	cttcttacttattcaagaagccatgaacatcatgtttatcctagctatgtattaccattg	1155
Db	31267	ATTCCTCTTATTTATAAAGCGATGAATTTATGTGTTCAATTTGCGATGTACTACCATTTG	31326
Qy	1156	ccacacactaggaatggaatgggctggatcgtggtccatggctgggcatcctctgcttcgacag	1215
Db	31327	TCACACCTCGGGCTGGATGGGTTGGATTTGGTCAATGGCTGGTATTTATGCTTTGATAG	31386
Qy	1216	gtctgcgcaattgtacgtatcatcatgaacgagggtcttaagaccgcacatttctgtogac	1275
Db	31387	ATTTCTGCAGGATTTGTAGAATTAATCATGAATGGTGGCTTGAAAACTGCTACTTCAGTAC	31446

Qy	1276	cacagatgattcttaacg	ttaacaagatctctgt	ccaagaagcctaag	tctcttcaag	tatca	1335
Db	31447	CACGTATGATTCTTA	TGTTATTAATAATTC	AGTTTCAAGTAA	AAAAAACCAAG	TTTTTTCACGATACCA	31506
Qy	1336	agtggaacattgcctta	tatgtacttctcttca	caaaatcagctggtt	ctctcacagttt	1395	
Db	31507	AGTAGAGAGCTTTC	GCATACATGATTTCT	TATACACCAAAAAG	TGCTATGAGTTT	31566	
Qy	1396	tcaattctatcccttca	cgcttctatcagaag	gcacagagatcttca	caacccagatca	1455	
Db	31567	CCAATCACATCCAT	TATACAGTATTATTC	GACGACACCGTGAT	CCAAACAATCCAGATCA	31626	
Qy	1456	actaactatgtacgtc	aaagctaaagggcaat	tacgagagtagtctt	cttcaagaagtctt	1515	
Db	31627	ATTGACGATGTACG	TAAAGGCAAAATAA	AGGTATCACTCGCAG	TTTTTGTATCGAAAGTTCT	31686	
Qy	1516	aagcgctccaacatac	cggttgattgacgaat	tctttagagagaccat	atgcgctaac	1575	
Db	31687	AGTGCTCCAAATCA	TACTGTTGATTGTA	ATTAATTCCTTGAAG	CCCAATATGGTGTAAAC	31746	
Qy	1576	tgtcctcacattgccc	aacttaagagaaatc	tagtagagtagctg	cgggcctcgcggt	1635	
Db	31747	GGTTCACATATCG	TAAAGTAAAGAAAT	CTGTGTAGCTGTAG	CCGCTGGTTTGGGTGT	31806	
Qy	1636	ggcagcactctacccc	atttctgtagaatgc	cttagattgcttagac	tatgatacactgata	1695	
Db	31807	TGCGGCTATTATC	CGCATTTTGTGCAAT	TGTTACGGTTACCAT	CTACTGATCAACTTCA	31866	
Qy	1696	gcacagttctactgat	cgtcaacgaccttag	tcaccttaagtggttc	gaaacgagct	1755	
Db	31867	GCATAAATTTTACT	GGATTGTTAATGACCT	ATCCCATTTGAAATG	GCTTTGAAATGAATTT	31926	
Qy	1756	acaatggcttaaggga	aaactcttggaagctc	tctgtcatctacactc	aggtggtcatcagtgga	1815	
Db	31927	GCAATGGTTAAAG	GAGAAAGTTGTAAG	CTCAGTCAATATACT	TGTTCCAGTGTGA	31986	
Qy	1816	ggatacaactcagatg	agtcacactaagggtt	ctcgatgacaaggaaga	aatctgaaatcac	1875	
Db	31987	GGACACAATTTCA	GATGAGAGTACAAA	AGGTTTGTATGATATA	AAAGAAACGCAAAATCAC	32046	
Qy	1876	cgtagaatgccttaac	agagagccagacctc	aaagagctagtgagat	cagagatcaaat	1935	
Db	32047	TGTTGAATGCTCAT	ATAAGACCTGATT	TGAAGAACTAGTGC	CTCGGAAATAAAACT	32106	
Qy	1936	gtcagaactcagaga	acaacatacacttct	tactcatgcggaccag	cgcacttccaatga	1995	
Db	32107	CTCAGAACCTAGAG	AAATAATAATATT	TACCTTTTATTCCTCG	CGGCCAGCAAGCTTTAAGCA	32166	
Qy	1996	cgaacttaggaatgc	agttgtacaaggtat	cgaattctagctgaag	atagatgtcgaaact	2055	
Db	32167	CGATTTTGAATGCA	GTGGTCCAGGTATAG	ACTCTCCTCTTGAAGAT	TTCAGCTTTGAACT	32226	
Qy	2056	agagagagagagttt	tacttggaagagctc	aaagct	2091		
Db	32227	AGAAGAAGAAGTTT	TACATGTAAGCCCT	TGTT	32262		
RESULT	4						
LOCUS	A86082		694 bp	DNA	PAT	21-JAN-2000	
DEFINITION			Sequence: 741 from Patent	EP0866129.			
ACCESSION	A86082						
VERSION	A86082.1		GI:6734681				
KEYWORDS							
SOURCE							
ORGANISM							
REFERENCE							
AUTHORS							
TITLE							
JOURNAL							

RESULT	4
A86082	
LOCUS	A86082 694 bp DNA
DEFINITION	Sequence.741 from Patent EP0866129.
ACCESSION	A86082
VERSION	A86082.1 GI:6734681
KEYWORDS	
SOURCE	Eremothecium gossypii.
ORGANISM	Eremothecium gossypii Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes; Saccharomycetales; Eremotheciaceae; Eremothecium. 1 (bases 1 to 694) Mohr, C. and Knechtle, P. Genomic DNA sequences of <i>Ashbya gossypii</i> and uses thereof Patent: EP 0866129-A 741 23-SEP-1998;
REFERENCE	
AUTHORS	
TITLE	
JOURNAL	

	FEATURES	CIBA GEIGY AG (CH)	Location/Qualifiers 1..694 /organism="Eremothecium gossypii" /db_xref="taxon:33169"		
	SOURCE	source			
	BASE COUNT	164 a	154 c	189 g	187 t
	ORIGIN				
		Query Match	7.1%; Score 147.6; DB 6;	Length 694;	
		Best Local Similarity	54.6%; Pred. No. 4.7e-32;		
		Matches 344; Conservative	0; Mismatches 274;	Indels 12; Gaps 2;	
	Qy	1006 ctacttcagatggggaatagtgcacaaattcttattgtccatcatcattttccagtccga	1065		
	Db	76 CTATTTTAAATGGGCTGTGTCACACGGTTTTTGCCCGGCTTCTTGCTCCTGCATAGTA	135		
	Qy	1066 gaaggcttcagaaccgccgagggttaatactcttcttacttatccacaagccatgaacat	1125		
	Db	136 AAGATTCAATCGGCCAAGAAGACGTACGATACTCTTGTACTGACACAAGCTCTTCAACAT	195		
	Qy	1126 catgtttatatcatgatgtattaccattgccacacactagatggatggctcgatctg	1185		
	Db	196 TGCTTTTATTGATGATCATATATGCACATCAAACGCTGGGATGGCACGGCTGGGTCTG	255		
	Qy	1186 gtccatggctggcactctctgttcgacaggttctgcgaattgttacgtatcatacatgaa	1245		
	Db	256 GTCGATGGTTGCCATCTACTGCTTCGAGCGTGTGCCCGGATAGCTCGCATTTGACTTGC	315		
	Qy	1246 cggagctttaagaccgcgcacctttgtcacacacagatgattcaacgtttacaagatctc	1305		
	Db	316 TGGAGGCATCAAGAAGCCACATTT---AACAGATGTTGGGGATCGCGTCAAGATGAC	372		
	Qy	1306 tgtcaagaagcctaagtcttccaagtcaagtggagcatttgctctatgtactttct	1365		
	Db	373 AGTGGAGAAGCCAAGACATTTCAAAATATTACC CGGGGCTTATGTTTTCGTTATTATAT	432		
	Qy	1366 ttaccacaaatcagcctggtttctaacgttttcaatctccatcccttcacagtcctcatcaga	1425		
	Db	433 TAGTGGGAAGGATGCTTGGTGTCTATCCATTCAGTCGCAACCGTTACCGCTCTTAATAC	492		
	Qy	1426 auggcacagagatcctaacaaccagatcaactaactatgacgtcaagctcaacaagg	1485		
	Db	493 ACCCAAGATCGATGGCGACAACTGGTG-----ATTATTTCAAAAGTGCACAAAGG	543		
	Qy	1486 cattacagagtagtactcttagcaaatcttaagcgtccaaaaccataccgttgattgcaa	1545		
	Db	544 CGTGACGACAGCTGCTAAACAGGATCTTTCTATCCGGGAAAGAGTCCATCGAATACAA	603		
	Qy	1546 gattttcttagaggaccatataggcgtaactgtcccctcacattgccccaaacttaagagaaa	1605		
	Db	604 GGTCGCTTCTAGAAGGCCCTATGGAACACCATTCGCGGGCTGTGCTCTCGACCGCG	663		
	Qy	1606 tcctagtagtagtgcggtgcgcctgcgct	1635		
	Db	664 CTACGTGGCGCCAGCGCAGGCTTGGCGT	693		
	RESULT	5			
	LOCUS	AR155575	694 bp	DNA	PAT
	DEFINITION	Sequence 741 from patent US 6239264.			08-AUG-2001
	ACCESSION	AR155575			
	VERSION	AR155575.1			GI:15123628
	KEYWORDS	Unknown.			
	SOURCE	Unknown.			
	ORGANISM	Unclassified.			
	REFERENCE	1 (bases 1 to 694)			
	AUTHORS	Philippsen,P., Pohlmann,R., Steiner-Lange,S., Mohr,C., Wendland,J., Knechtie,P. and Rebischung,C.			
	TITLE	Genomic DNA sequences of ashbya gossypii and uses thereof			
	JOURNAL	Patent: US 6239264-A 741 29-MAY-2001;			

[illegible]

LOCUS	D83181	3084 bp	DNA	PLN	06-FEB-1999
DEFINITION	Candida albicans DNA for CFL1, complete cds.				
ACCESSION	D83181				
VERSION	D83181.1	GI:1783272			
KEYWORDS	CFL1				
SOURCE	Candida albicans DNA.				
ORGANISM	Candida albicans				
	Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes; Saccharomycetales; mitosporic Saccharomycetales; Candida.				
REFERENCE	1. (bases 1 to 3084)				
AUTHORS	Okabe,T.				
TITLE	Direct Submission				
JOURNAL	Submitted (19-JAN-1996) to the DDBJ/EMBL/GenBank databases. Toshiko Okabe, Yokohama City University, School of Medicine, Department of Hygiene; 3-9 Fukuura, Kanazawa-Ku, Yokohama, Kanagawa 236-0004, Japan (E-mail:toshiko@med.yokohama-cu.ac.jp, Tel:81-45-787-2606, Fax:81-45-786-5219)				
REFERENCE	2. (bases 1 to 3084)				
AUTHORS	Yamada-Okabe,T.				
TITLE	Isolation of mRNA capping enzyme from Candida albicans, similar but distinct genome organization between Candida albicans and Saccharomycetaceae				
JOURNAL	Unpublished (1996)				
REFERENCE	3. (sites)				
AUTHORS	Yamada-Okabe,T., Shimizu,O., Doi,R., Mizumoto,K., Arisawa,M. and Yamada-Okabe,H.				
TITLE	Isolation of the mRNA-capping enzyme and ferric-reductase-related genes from Candida albicans				
JOURNAL	Microbiology 142 (pt 9), 2515-2523 (1996)				
MEDLINE	96425877				
FEATURES	Location/Qualifiers				
source	1. .3084				
	/organism="Candida albicans"				
	/db_xref="taxon:5476"				
	401..2410				
	/codon_start=1				
	/transl_table=12				
	/product="CFL1"				
	/protein_id="BAAL1834.1"				
	/db_xref="GI:1783273"				
	/translation="MTSEKPHAKYDIQAEFTNGTTEYAKMTTKSSGSKTSTASKS SKTSGSNASKSTNAHGSNSTSSSSSSKSGNSGTSTTETITPLLDYKFFT PKYDAQSMNNENLSTNYSGLLGYWAGLAIAPANIKMKFPLTNLNSIGNL FRKHLDPATFERKKQAERFSGVGFDFGLIPTRLETIIWFEVLTLGLFSACHIHV KDNPOYATNVAIEHTFESVDKATGYKKNMKRDPMTWGTVCIGGFLPQAMLFRH WISRVDFLLIHVILVVFVVGYYHLESQGYDFMMAAATAVADFRRVLRIGRIFFGA RKATYIKGDDTLKIEVPKPKYKWSVAGGHAFTHFLKPTLFQSHPTFTFTTSGNKI RVATKINGTTSNIAKYLSPCLNTATIRLVVEGPGYESSAGRNCKNVFVAGGDKI PGYISCVDLAKSKNSKIKLWIRHWKLSWFTBELEYLKKTNQSTIYVTPQDCC SGLCEPHDVFEKSKDERKDSVSSQYSLISNLIKQSLHVEFTEGSRPDI5TQVEQEV QADGAIGFVTCGHPAMVDLURFATONLNVSKRIVETHQLQLTWA"				
BASE COUNT	885 a 583 c 601 g 1015 t				
ORIGIN					

	Query Match	4.88;	Score 99.4;	DB 8;	Length 3084;
	Best Local Similarity 47.1%;	Pred. No. 9.7e-18;			
	Matches 344;	Conservative 0;	Mismatches 381;	Indels 6;	Gaps 1;
QY	648	ttcaactcgaggcaaggaactcgtagcttctattctgttcattctgactattctctcac	707		
Db	972	TTCCAACGAGATTGAAACAATAATTGTTGTTATTTTGTGCTTTCAGGGGCTCTCTTCA	1031		
QY	708	tctcttcggacatacaatcaagttgcacatctctacgatagacatagatggagaagat	767		
Db	1032	GTGCCTTAGCATATTCATCAGCTCAAGATAATCAACAATATGCTACTAAGAATGCTGAAT	1091		
QY	768	caatggcattcgtctcaagccgctgctgaacttgatggcaatcgctcttttcccgctggatgt	827		
Db	1092	TGGGTGATTTTGTATTGCCGACAGGACTGGTATTTTATAGAACATTTTGTATTGCTGTTGTA	1151		
QY	828	accttttcggtatcccggaacaaccccttcattcccaatcacccggagattgagcttttagtactt	887		

Db 1152 TCCTGTTGGTGGTGTAAACAACTCTTACAAAGCTAACTGGTGGGATTTTGCTACTT 1211
Qy 888 tcaacttttaccacaaatggtcagacatcgtctgtctcatgtagcgcgtccatccaa 947
Db 1212 TTATCATGTATCAGAGATGGATTTCCCGTGTGTGATGTTTATTTGATCATCGTTCATGCAA 1271
Qy 948 tcgttatgaccgttcaggagtaaacagagagtagtaccagctctcttaaggaatctc 1007
Db 1272 TCACATTTTCGTGCTCACAAGCAACGGGCAATATACAAAACAGATGAACCTGACT 1331
Qy 1008 acttcagatggggaatagtagccacaattcttatgcccatacatctttccagtcgcgaga 1067
Db 1332 TCATGATTTGGGGTACAGTTCTACAATATGAGGAGATTTATTCTTTTCAGGCCATGT 1391
Qy 1068 aggtcttcaggaacagaggtatgaaatctcttcttcttattcacaagaacccatgaacatca 1127
Db 1392 TATTTTTCAGAAAGAAATGTTATGAAGTCTTCTTTTTCATCCATATGTTTGGTGTGTT 1451
Qy 1128 tggttatcatagctatgattaccattgccacacactaggatggtggctggatcgtgt 1187
Db 1452 TCCTTTGTTGGGGGATACATCAATTTGGAAGTCAAGGATATGCGGATTCATGCTGGG 1511
Qy 1188 ccagtggtggaatcctctgctgcagacaggttctgcggaattgtacgtatcatcatgaacg 1247
Db 1512 CGCCATTGCTGTTGGGCTTTTGATCGTGTCTCAGATTAGGAGAAATTTTCTTTTCG 1571
Qy 1248 gagggtcttaagaccgcactttgtgcacacagatgattctaacgttatcaagatctctg 1307
Db 1572 GTGCGAGAAAAGCAACCGTCTCTATTAAAGGTGAGCATAC-----TTGAAGATCGAGG 1625
Qy 1308 tcaagaagcctaagttcttcaagtagcatcaagtgagagcattgcttatgattcttctt 1367
Db 1626 TTCCCAACCTAAATATTGGAAGTCGGTAGCAGGTGGACATGCTTTATTCAATTCCTGA 1685
Qy 1368 caccaaaatca 1378
Db 1686 AACCAACCTTA 1696

RESULT 8
CAL387722 3615 bp DNA PLN 29-APR-2000
LOCUS Candida albicans cf11 gene for ferric reductase.
DEFINITION
ACCESSION AJ387722
VERSION AJ387722.1 GI:5139349
KEYWORDS cf11 gene; ferric reductase.
SOURCE Candida albicans.
ORGANISM Candida albicans.
Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
Saccharomycetales; mitosporic Saccharomycetales; Candida.
REFERENCE 1 (bases 1 to 3615)
AUTHORS Hammacott,J.E., Williams,P.H. and Cashmore,A.M.
TITLE Candida albicans CF11 encodes a functional ferric reductase
activity that can rescue a Saccharomyces cerevisiae frel mutant
JOURNAL Microbiology 146 (Pt 4), 869-876 (2000)
MEDLINE 20244631
REFERENCE 2 (bases 1 to 3615)
AUTHORS Hammacott,J.E.
TITLE Direct Submission
JOURNAL Submitted (18-JUN-1999) Hammacott J.E., Genetics, University of
Leicester, University Road, Leicester, LE1 7RH, UNITED KINGDOM
FEATURES
source
1..3615
/organism="Candida albicans"
/db_xref="taxon:5476"
CAAT_signal 920..923
gene 920..3346
TATA_signal 1041..1045
CDS 1064..3346

/gene="cf11"
/note="Protein sequence is in conflict with the conceptual translation"
/codon_start=1
/transl_table=12
/codon=(seq:"cag",aa:Ser)
/product="ferric reductase"
/protein_id="CAB45649.1"
/db_xref="GI:5139350"
/db_xref="SPTREMBL:Q9Y861"
/translation="MKIQQLIVFLFAVVLIDATPKRYSELDIVMSTCTTPIGKYGTV
CTSTGKRSTNWCYKTDAGFTISDCLVRGNNTNIISKFTSCNNTEKSFHAKYD
KISAEFTNGTYAKMTTKSSGSKTSASASKSTGSSNASKSTNAHGSNSTSS
TSSSSKSGKNSCTSTETITPTLLIDYKFTPTPKAYOMNNNNLSINVGSLILG
YWAGILAIAPANNIKMFPSTLNLSSISNLPKHLFLPATPKKKAQSFSLGVG
FFDGLIPRLTIIIVIFVLTLGLFSALHIIHVADNPQYATKNAELGHLIADRTGILG
TFLPLILIFGGRNFILOWJTGWDFATFIYHWRWISRVVLLIILVHAITFVSVDKATG
KYLKMRKDFMINTGTVTCGGTFLFSAMLFRRKCYEFVLIHIVFVFPKYWK
LESQGYGDFMAAIAVWAFDRVRLGRFFFGARKATVSIKGDITLIEVKPLPYWKS
VAGHAFTHFLKPTLFLQSHPTFTTSTESNDKIVLYAKIKNGITSNIAKSIKLIWII
ATISLVGEPYGEPSAGRNCKNVFVAGNGIPCIYSECVDLAKKSKNSIKLIIWII
RHWKLSWFTSELEYLAKTNVQSTIYVTPQDCSGLECFEHDVSEFKSDKSDVESS
QYLSINIKOGLSHVEFTGHRPDIISTQVEVQKQADGAIGFVTCGHPANVDELPAVT
SNLNVSKHRVYEHQLSWA"
BASE COUNT 1183 a 596 c 612 g 1224 t
ORIGIN

Query Match 4.7%; Score 97.8; DB 8; Length 3615;
Best Local Similarity 46.9%; Pred. NO. 3e-17;
Matches 343; Conservative 0; Mismatches 382; Indels 6; Gaps 1;
Qy 648 ttacaactcaggcgaaggactgtagttcttattcttctgcttctgacttatctctc 707
Db 1908 TTCCAACAGATATGAGAAACAATATGTTGTTATTTGTTGCTGTGAGGGTCTCTCA 1967
Qy 708 tctcttcggacataacatcaagttgcccacatccctacagatagactagatgagaagt 767
Db 1968 GTGCTTACATATTCATCAGTCAAGATATCCACAATATGCTACTAGATGCTGAAT 2027
Qy 768 caatggcattcgtctcagccgctgctgacttgatggcaatcgctctcttcccccggtgt 827
Db 2028 TGGGTCATTTGATGTCGACAGGACTGGTATTTAGGAACAATTTTGTATTCGTTGTTGA 2087
Qy 828 acctttcggatccggaaacaccccttcaccccaatcaccggattgagcttagtactt 887
Db 2088 TCTTGTGTTGGTGTCTAACAACCTTCTTACAATGGCTAACTGTTGGGATTTTGTCTACT 2147
Qy 888 tcaacttttaccacaaatggtcagcatcgtctgcttcatgtagcgcgtccatccaa 947
Db 2148 TTATCATGTATCAGATGAGATTTCCCGTGTGTGATGTTTATTTGATCATCGTTCATGCAA 2207
Qy 948 tcgttatgaccgttcaggagtaaacagagagtagtattccagctctcttgaaggaattct 1007
Db 2208 TCACATTTTCGTGCTGACAAAGCAACGGGCAATATACAAAACAGATGAACCTGACT 2267
Qy 1008 acttcagatggggaatagtagccacaattctctatgcccatacatcttccagtcgcgaga 1067
Db 2268 TCATGATTTGGGGTACAGTTTCTACAATATGTGGAGGATTTATTTCTTCAGGCCATGT 2327
Qy 1068 aggtcttcaggaacagaggttatgaaatctcttcttcttattcacaagaacccatgaacatca 1127
Db 2328 TATTTTTCAGAGAAATGTTATGAAGTCTCTTTTGTGATCCATATGTTTGGTGTGTTT 2387
Qy 1128 tggttatcatagctatgattaccattgccacacactaggatggtggctggatcgtgt 1187
Db 2388 TCTTTGTTGTTGGGGATACATCATTTTGGAAAGTCAAGGATATGCGGATTTTCATGTGG 2447
Qy 1188 ccagtggtggaatcctctgctgcagacaggttctgcggaattgtacgtatcaatgaacg 1247
Db 2448 CGCCATTGCTGTTGGGCTTTTGATGCTGTCGATAGGATAGGAGAAATTTTCTTTTCG 2507
Qy 1248 gagggtcttaagaccgcactttgtgcacacagatgattcttaacgttatcaagatctctc 1307

Db 2508 GTGCAAGAAAAGCAACCGTCTCTATTAAAGGTGACGATAC-----TTTGAAGATTGAGG 2561
Qy 1308 tcaagaagcctaagttctcaagtcaagtggagcatttgctctatatgtactttttt 1367
Db 2562 TTCCCAACACCTAAATATTGGAAGTGGTAGCAGGTGGACATGCTTTTATTTCATTCTTGA 2621
Qy 1368 caccacaaatca 1378
|||||
Db 2622 AACCAACCTTA 2632
|||||
RESULT 9
LOCUS A85974 711 bp DNA PAT 21-JAN-2000
DEFINITION Sequence 633 from Patent EP0866129.
ACCESSION A85974
VERSION A85974.1 GI:6734573
KEYWORDS
ORGANISM Eremothecium gossypii.
REFERENCE 1 (bases 1 to 711)
AUTHORS Mohr,C. and Knechtle,P.
TITLE Genomic DNA sequences of Ashbya gossypii and uses thereof
JOURNAL Patent: EP 0866129-A 633 23-SEP-1998;
CIBA GEIGY AG (CH)
FEATURES
source
1. .711
/organism="Eremothecium gossypii"
/db_xref="taxon:33169"
BASE COUNT 165 a 193 c 167 g 186 t
ORIGIN
Query Match 4.3%; Score 90.2; DB 6; Length 711;
Best Local Similarity 50.3%; Pred. No. 3.8e-15;
Matches 280; Conservative 0; Mismatches 268; Indels 9; Gaps 2;
Qy 1526 aaccataccgttgattgcaagattttcttagaggagaccataatggcgttaactgtccctcac 1585
Db 17 AAAGAGTCCATCGAATACAAAGTGCTTCTTAGAAGGGCCCTATGGAACACACCATTTCCGCGG 76
Qy 1586 attgccaaaactaagagaactctagtaggagtagctcgggcctcgcgtaggcagccatc 1645
Db 77 CTTCGCTCTCTGACGGCGTACGTGGGCCGCCAGCGAGGTCTTGGCGTAGCAGCGGTC 136
Qy 1646 taccctcattctgtagaatgccttagattgctgtagcactgatcaactgcagcacaagtcc 1705
Db 137 TACCCACACTTCGT-----CTCTCTGTGGACAAGGAAGCCAGTTACCCCATTCATTTC 190
Qy 1706 tactggatcgtaacgaccttagtcaccttaagtgggttcgaaaacgagactacaatggctt 1765
Db 191 TACTGGATTATAAATGACCTTTTCATATCTGCATTGGTTTTCGCATGAGCTGAGGTACCTG 250
Qy 1766 aaggagaaactcttgagtcctgtcatctacactgggtcatcagtgaggatcacaaac 1825
Db 251 GCGGACCGGAACCTCGACATCAAAATTTATTACAGGAGGACCAATGATGTCGGCTAAAGAA 310
Qy 1826 tcagatgagtcacactaagggtttcg---atgacaaggaagaatctgaaatcacccgtagaa 1882
Db 311 CTGACCCAGATGTTGCCGATTCCGCCCTCTCGGAAGTTTCGGGATTTCGCTGGATATCTGC 370
Qy 1883 tgccttaacaagagggcagacctcaagagcgtagtggatcagagatacaaatgtccagaa 1942
Db 371 AGGCTCTCTTCGCCCCAGATCTCAAAAGAGATCGTGGAAAGACAGATCCTCTCTCGTCT 430
Qy 1943 ctcgagaacaacaacatcactttctactcatcgagaccagcgactttcaatgacgacttt 2002
Db 431 AACACGGCACAGGACGTCACGCTTTATTAGCAGCGGCCCTTCGACCTTTTAAATGACCATTTTC 490
Qy 2003 aggaatgcagttgtacaagggtatcgattcttagtctgaagatagatgtcgaactagaggag 2062
Db 491 CGCTATGCTGTGAAATCTAGCATCACGGGCAAACTCCAGTGTGTGATGTCGACCTAGAGGAG 550
Qy 2063 gagagtttacttgta 2079
|||||
Db 551 GAAAGCTACACCTGGTA 567
|||||

Db 491 CGCTATGCTGTGAAATCTAGCATCACGGCAAACTCCAGTGTGTGATGTCGACCTAGAGGAG 550
Qy 2063 gagagtttacttgta 2079
|||||
Db 551 GAAAGCTACACCTGGTA 567
|||||
RESULT 10
LOCUS ARI55467 711 bp DNA PAT 08-AUG-2001
DEFINITION Sequence 633 from patent US 6239264.
ACCESSION ARI55467
VERSION ARI55467.1 GI:15123520
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 711)
AUTHORS Philippsen,P., Pohlmann,R., Steiner-Lange,S., Mohr,C., Wendland,J.,
Knechtle,P. and Reibischung,C.
TITLE Genomic DNA sequences of ashbya gossypii and uses thereof
JOURNAL Patent: US 6239264-A 633 29-MAY-2001;
FEATURES
source
1. .711
/organism="unknown"
BASE COUNT 165 a 193 c 167 g 186 t
ORIGIN
Query Match 4.3%; Score 90.2; DB 6; Length 711;
Best Local Similarity 50.3%; Pred. No. 3.8e-15;
Matches 280; Conservative 0; Mismatches 268; Indels 9; Gaps 2;
Qy 1526 aaccataccgttgattgcaagattttcttagaggagaccataatggcgttaactgtccctcac 1585
Db 17 AAAGAGTCCATCGAATACAAAGTGCTTCTTAGAAGGGCCCTATGGAACACACCATTTCCGCGG 76
Qy 1586 attgccaaaactaagagaactctagtaggagtagctcgggcctcgcgtaggcagccatc 1645
Db 77 CTTCGCTCTCTGACGGCGTACGTGGGCCGCCAGCGAGGTCTTGGCGTAGCAGCGGTC 136
Qy 1646 taccctcattctgtagaatgccttagattgctgtagcactgatcaactgcagcacaagtcc 1705
Db 137 TACCCACACTTCGT-----CTCTCTGTGGACAAGGAAGCCAGTTACCCCATTCATTTC 190
Qy 1706 tactggatcgtaacgaccttagtcaccttaagtgggttcgaaaacgagactacaatggctt 1765
Db 191 TACTGGATTATAAATGACCTTTTCATATCTGCATTGGTTTTCGCATGAGCTGAGGTACCTG 250
Qy 1766 aaggagaaactcttgagtcctgtcatctacactgggtcatcagtgaggatcacaaac 1825
Db 251 GCGGACCGGAACCTCGACATCAAAATTTATTACAGGAGGACCAATGATGTCGGCTAAAGAA 310
Qy 1826 tcagatgagtcacactaagggtttcg---atgacaaggaagaatctgaaatcacccgtagaa 1882
Db 311 CTGACCCAGATGTTGCCGATTCCGCCCTCTCGGAAGTTTCGGGATTTCGCTGGATATCTGC 370
Qy 1883 tgccttaacaagagggcagacctcaagagcgtagtggatcagagatacaaatgtccagaa 1942
Db 371 AGGCTCTCTTCGCCCCAGATCTCAAAAGAGATCGTGGAAAGACAGATCCTCTCTCGTCT 430
Qy 1943 ctcgagaacaacaacatcactttctactcatcgagaccagcgactttcaatgacgacttt 2002
Db 431 AACACGGCACAGGACGTCACGCTTTATTAGCAGCGGCCCTTCGACCTTTTAAATGACCATTTTC 490
Qy 2003 aggaatgcagttgtacaagggtatcgattcttagtctgaagatagatgtcgaactagaggag 2062
Db 491 CGCTATGCTGTGAAATCTAGCATCACGGGCAAACTCCAGTGTGTGATGTCGACCTAGAGGAG 550
Qy 2063 gagagtttacttgta 2079
|||||
Db 551 GAAAGCTACACCTGGTA 567
|||||

Qy	1943	ctcgagagaaacaaacataccttcttactcatgtaggacacagagacttccaatgaagaaattt	2062		
Db	431	AAGCAGGCACAGGACGCTACGTTTATTAGCAGCGGCCCTTCGACCTTTAATGACCAATTC	490		
Qy	2003	aggaatgcagttgtacaagatgcattcttagctgaagatagatgctgaactagagag	2062		
Db	491	CGTATGCTGTGAATCTAGCATCAGCGGCAAACTCCAGTGTGATGTCGACCTAGAGGAG	550		
Qy	2063	gagagtttacttcttggtta	2079		
Db	551	GAAGCTACACCTGGTA	567		
RESULT	12				
E31009					
LOCUS	E31009	83 bp	DNA		
DEFINITION	Method for transforming plant, the thus transformed plant and genome thereof.				
ACCESSION	E31009				
VERSION	E31009.1	GI:13021495			
KEYWORDS	JP 1999266876-A/4.				
SOURCE	unidentified.				
ORGANISM	unidentified.				
REFERENCE	1 (bases 1 to 83)				
AUTHORS	Satoshi,M.H.O.O. and Nakanishi,H.Y.				
TITLE	Method for transforming plant, the thus transformed plant and genome thereof				
JOURNAL	Patent: JP 1999266876-A 4 05-OCT-1999;				
COMMENT	SCIENCE & TECH AGENCY				
	OS	Unidentified			
	PN	JP 1999266876-A/4			
	PD	05-OCT-1999			
	PF	24-MAR-1998 JP 1998096637			
	PI	SATOSHI MORI,HIROYUKI OKI,NOBUHITO NAKANISHI, PI			
	YAMAGUCHI				
	PC	C12N15/09,A01H5/00,C12N5/10,C12P19/34/((C12N15/09,C12R1:645),			
	PC	(C12N5/10,C12R1:91),C12N15/00,C12N5/00,C12N15/00,C12R1:645),			
	PC	(C12N5/00,C12R1:91)			
	CC	Strandedness: Single;			
	CC	Topology: Linear;			
	FH	Key			
	FT	source			
	FT	Location/Qualifiers			
FEATURES	Location/Qualifiers				
source	1..83				
	/organism="unidentified"				
BASE COUNT	17 a	26 c	13 g		
ORIGIN	27 t				
Query Match	4.0%; Score 83; DB 6; Length 83;				
Best Local Similarity	100.0%; Pred. No. 3.4e-13;				
Matches	83; Conservative	0; Mismatches	0; Indels		
Qy	1	gaattctctagatccacatggttagaacagagtcctttctgcctcttcacatcttt	60		
Db	1	GAATCTCTAGATCCACCATGTTAGAACACAGAGTCTTTCTGCCTCTCATCTCTTT	60		
Qy	61	cttcgctacagtcacatcgagcg	83		
Db	61	CTTCGCTACAGTCCATCGAGCG	83		
RESULT	13				
E31010					
LOCUS	E31010	83 bp	DNA		
DEFINITION	Method for transforming plant, the thus transformed plant and genome thereof.				
ACCESSION	E31010				
VERSION	E31010.1	GI:13021496			

Oy 272 agtatcaaggtttacacactgga 294
|||||
Db 23 AGTATCAAGGTTTACACACTGGA 1

Search completed: May 14, 2002, 07:24:06
Job time: 11053 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 14, 2002, 04:19:48 ; Search time 1786.89 seconds
(without alignments)
12580.611 Million cell updates/sec

Title: US-09-646-825-1

Perfect score: 2092

Sequence: 1 gaattcttagactccacca.....cttggtaagactcaagctt 2092

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 11351937 seqs, 5372889281 residues

Total number of hits satisfying chosen parameters: 22703874

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1: em_estfun:*

2: em_esthum:*

3: em_estin:*

4: em_estom:*

5: em_estpl:*

6: em_estba:*

7: em_estro:*

8: em_estov:*

9: em_htc:*

10: gb_estl:*

11: gb_est2:*

12: gb_htc:*

13: gb_gss:*

14: em_gss_fun:*

15: em_gss_hum:*

16: em_gss_inv:*

17: em_gss_pln:*

18: em_gss_pro:*

19: em_gss_rod:*

20: em_gss_vrt:*

21: em_gss_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result NO.	Score	Query Match	Length	ID	Description
c 1	187.4	9.0	962	13 CNS060YA	AL408680 T7 end of
c 2	149	7.1	515	13 A2931275	A2931275 474.dhz64
c 3	71.4	3.4	966	13 CNS062HJ	AL422333 T3 end of
c 4	69	3.3	968	13 CNS06HK9	AL399103 T7 end of
c 5	53.4	2.6	1015	13 CNS06VLD	AL421175 T7 end of
c 6	50.6	2.4	958	13 CNS06WLF	AL417865 T7 end of
c 7	48.2	2.3	851	13 CNS079Q2	AL435800 T3 end of
c 8	47.4	2.3	987	13 CNS06KPL	AL420031 T7 end of
c 9	44.2	2.1	1101	13 CNS01705	AL107423 Drosophil
c 10	44	2.1	805	13 A0858562	A0858562 nbe00020
c 11	42.8	2.0	888	13 A2528430	A2528430 ENT024TR
c 12	42.8	2.0	949	13 CNS06NMI	AL406960 T3 end of

13	42.8	2.0	1101	13	CNS0039G	AL063921 Drosophil
14	41	2.0	1043	13	CNS07DHO	AL440482 T3 end of
15	41	2.0	2149	11	EG714799	BG714799 602677117
16	40.8	2.0	391	10	AV677876	AV677876 602677117
17	40.6	1.9	843	13	AZ551618	AZ551618 ENT0V54TR
18	40.6	1.9	905	13	AZ550256	AZ550256 ENTEV58TR
19	40.2	1.9	908	13	AZ548467	AZ548467 ENTEK30TR
c 20	40	1.9	891	13	AZ683582	AZ683582 ENTKK47TR
c 21	39.6	1.9	300	10	AI899200	AI899200 EST268643
c 22	39.6	1.9	416	10	AI483599	AI483599 EST249449
c 23	39.6	1.9	487	13	AZ111934	AZ111934 RPCI-23-4
c 24	38.6	1.9	880	13	AZ669474	AZ669474 ENTY88TR
c 25	38.8	1.9	320	10	BB139106	BB139106 BB139106
c 26	38.8	1.9	974	13	CNS0605J	AL407645 T3 end of
c 27	38.8	1.9	1101	13	CNS00KDX	AL077856 Drosophil
c 28	38.6	1.8	633	10	AU214668	AU214668 AU214668
c 29	38.6	1.8	816	13	AZ535744	AZ535744 ENT0C25TR
c 30	38.2	1.8	1630	12	AK009327	AK009327 Mus muscu
c 31	37.8	1.8	328	13	AZ923069	AZ923069 4908.9e87
c 32	37.6	1.8	287	10	BB370647	BB370647 BB370647
c 33	37.6	1.8	290	11	BG937781	BG937781 1A006H02
c 34	37.6	1.8	1272	11	BG394360	BG394360 602456372
c 35	37.4	1.8	369	13	AQ152254	AQ152254 HS.2223_B
c 36	37.4	1.8	505	10	AI333640	AI333640 qp91d12.x
c 37	37.4	1.8	549	10	AL500267	AL500267 AL500267
c 38	37.2	1.8	584	11	BG463815	BG463815 EM1_51_C0
c 39	37.2	1.8	621	10	AU203061	AU203061 AU203061
c 40	37.2	1.8	716	10	AU213521	AU213521 AU213521
c 41	37	1.8	519	13	AZ803420	AZ803420 2M0063L15
c 42	37	1.8	668	13	AQ454443	AQ454443 HS.5195_A
c 43	37	1.8	944	12	AK015703	AK015703 Mus muscu
c 44	36.6	1.7	849	13	AZ546009	AZ546009 ENT0F53TF
c 45	36.4	1.7	321	11	BG224888	BG224888 kp51908.y

ALIGNMENTS

RESULT 1
CNS060YA/c

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

REFERENCE

AUTHORS

CNS060YA 962 bp DNA GSS 05-JUL-2001
T7 end of clone AVOAA010B01 of library AVOAA from strain CBS 379 of
Saccharomyces exiguus, genomic survey sequence.

AL408680

AL408680.1 GI:12175485

GSS.

Saccharomyces exiguus.

Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;

Saccharomycetales; Saccharomycetaceae; Saccharomycetes.

1 (bases 1 to 962)

Bon,E., Neuveglise,C., Lepingle,A., Wincker,P., Artiguenave,F.,

Gaillardin,C. and Casaregola,S.

Genomic exploration of the hemiascomycetous yeasts: 6.

Saccharomyces exiguus

FEBS Lett. 487 (1), 42-46 (2000)

20584716

2 (bases 1 to 962)

Souci,J.L., Aigle,M., Artiguenave,F., Blandin,G.,

Bolotin-Fukuhara,M., Bon,E., Brottier,P., Casaregola,S.,

de-Montigny,J., Dujon,B., Durrens,P., Lepingle,A., Llorente,B.,

Malpertuy,A., Neuveglise,C., Ozier-Kalogeropoulos,O., Potier,S.,

Saurin,W., Tekala,F., Toffano-Nioche,C., Wesolowski-Louvel,M.,

Wincker,P. and Weissenbach,J.

Genomic exploration of the hemiascomycetous yeasts: 1. A set of

yeast species for molecular evolution studies

FEBS Lett. 487 (1), 3-12 (2000)

20584711

3 (bases 1 to 962)

Genoscope.

Direct Submission

Submitted (07-SEP-2000) Genoscope - Centre National de Sequencage,

2 rue Gaston Cremieux, CP 5706, 91057 EVRY cedex, FRANCE. (E-mail :

seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)
This GSS is part of a random genomic sequencing program of thirteen yeast species: *Saccharomyces bayanus* var. *uvarum*, *Saccharomyces exiguus*, *Saccharomyces servazzii*, *Zygosaccharomyces rouxii*, *Saccharomyces kluyveri*, *Kluyveromyces thermotolerans*, *Kluyveromyces lactis* var. *lactis*, *Kluyveromyces marxianus* var. *marxianus*, *Pichia angusta*, *Debaryomyces hansenii* var. *hansenii*, *Pichia sorbitophila*, *Candida tropicalis* and *Yarrowia lipolytica*. Genomic inserts of 3 to 5 kb were prepared and both extremities were sequenced. See keywords for description of this sequence and for the sequence of the other extremity of this insert.

FEATURES

```

1. 1962
/organism="Saccharomyces exiguus"
/strain="CBS 379"
/db_xref="taxon:34358"
/clone="AV0AA010B01"
/clone_lib="AV0AA"
/notes_end="T7"

```

BASE COUNT	355 a	165 c	304 t	2 others
ORIGIN				

Db	342	ACAGTAATGATGTCGGGTATTATGTGCCACATAGTCAGAAAATAATTATAGAGATAGATTCTAT	283
Qy	1091	gaatctcttactattatcacaagaagccatgaacatcatgtttatcatagctatgtattac	1150
Db	282	GAATGTTTCTATTCTTTTCATAAAGCGATGAATATATCGTCATCATGGCATATATTTC	223
Qy	1151	cattgccacacatagatggaatggctcggtatctgttcacatgctgagtcctcigtctc	1210
Db	222	CATATTAAATCATTTGGTGTGGCTAAATTTGGGTTTGGGCACCTTGTGGTATCTGGGGCGATG	163
Qy	1211	gacaggttctgccgaattgtcgtatcatcatgaacgaggtctcaagaccgccactttg	1270
Db	162	GATAGGGTATTGAGGATGTTTGGGTATTATTAAATGGTGGGATTCATAATGCTACATT-	102
Qy	1271	tcgaccacagatgattctcaagcttatcaagatcctctgccaagaagcctaagttctcaag	1330
Db	103	--AACAGATTGTTACAAATGGTGTAATTAGCTTCTCTATACCTAAACCAGGATTCITTA	46
Qy	1331	tatcaagtggagcatttgcctatatgtacttt	1363
Db	45	TATCAACAGGATGTTTGTATGATGATATTCT	13

	Query Match	7.18;	Score 149;	DB 13;	Length 515;
	Best Local Similarity	56.4%;	pred. No. 2e-30;		
	Matches 278;	Conservative 0;	Mismatches 215;	Indels 0;	Gaps 0
Qy	720	ataacatcaagttgccacatccttacgatagccttagatggagaagatccaatggcaattcg	779		
		9	ATAATATTACTTTACCACATCCCTATATGTCATCTAGATGGTATGCAAAATTTATATGTGG	68	
Db					
Qy	780	tctcaagccgcgtgcgaactgtagtgcgaatcgctctcttttcccgcgggtgcacctttttgggta	839		
Db	69	TTAGCTTATAGAACAGATTTAATGGCAATGGGCAATATTTCCATTAAATTTCTTTTGGCA	128		

BASE COUNT 241 a 263 c 273 g 232 t 6 others
ORIGIN

Query Match 2.6%; Score 53.4; DB 13; Length 1015;
Best Local Similarity 46.9%; Pred. No. 0.00091;
Matches 196; Conservative 2; Mismatches 214; Indels 6; Gaps 1;

QY 148 gcaagaacttctactcaagaacatacaattggctcggaagcgtcaactgcagtcgctta 207
DB 445 GGCATACCGTTTCGGTGGCGCAACGTGAACCTGGCTCGGCACCATCAACCAATTGTATCGC 504
QY 208 tgagaactccaataatacaagaactctggaactccgttgatgaacttgccagcagaatg 267
DB 505 AAACAACACCAAGACCTGGCTCAGGACCAACCGGGTTCGGCAGCTTCGGCGGAGGTG 564
QY 268 ctcaagtatcaaggtt-----tacacactggaggacatgaagaacatacttaactaatgc 321
DB 565 CTACGAGAAAGGGGATTTCAACTACACTCTGGCGGACATGCGCGCTTCTACGCTAACGG 624
QY 322 aagtaactactctcgctctctgagaataccgataagaagacagttgtttcaaacacgtt 381
DB 625 CACAGGTACCTCAGAGACCCACCACAGCGGACCTGTCTAGCCCGGTTTTCACACCGCT 684
QY 382 gatggcaaatgagcggcctatactactactactatgaggaactatgggatccactgaa 441
DB 685 TCGSTGANACCAAGCGAGTTCGATGCTGACTACAAGAGATGAAGACCTCAGCTTCTC 744
QY 442 ttgatgcatatcaatggtgcgcatggggcctctctctctctctctctctctctctcttac 501
DB 745 TGTGACGCGCAGTCAAGTGGTTCGGGTGGCGCTGGTGTGTTTCTGGCGGACATATTGC 804
QY 502 cgcgcgaactatctgaacatctcaaacgcgtattcggcaagaacattatggcaaat 559
DB 805 CGTACCACGTTGTCAATTTCAACGAAAAGTTTTRTGGGCTGAGCTCACCAGCAACT 862

RESULT 6

CNS06W1F 958 bp DNA GSS 06-JUL-2001
LOCUS T7 end of clone AX0AA027B12 of library AX0AA from strain CBS 7064
DEFINITION of Pichia farinosa, genomic survey sequence.
ACCESSION AL417865
VERSION AL417865.1 GI:12200119
KEYWORDS GSS.
SOURCE Pichia farinosa.
ORGANISM Pichia farinosa.
REFERENCE Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
Saccharomycetales; Saccharomycetaceae; Pichia.
AUTHORS 1 (bases 1 to 958)
de Montigny, J., Spohner, C., Souciet, J., Tekala, F., Dujon, B.,
Wincker, P., Artiguenave, F. and Potier, S.
TITLE Genomic exploration of the hemiascomycetous yeasts: 15. Pichia
sorbitophila
JOURNAL FEBS Lett. 487 (1), 87-90 (2000)
MEDLINE 20584725
REFERENCE 2 (bases 1 to 958)
Souciet, J., Aigle, M., Artiguenave, F., Blandin, G.,
Bollotin-Fukuhara, M., Bon, E., Broutier, P., Casaregola, S.,
de Montigny, J., Dujon, B., Durrens, P., Lepingle, A., Llorente, B.,
Malpertuy, A., Neuveglise, C., Ozier-Kalogeropoulos, O., Potier, S.,
Saurin, W., Tekala, F., Toffano-Nioche, C., Wesolowski-Louvel, M.,
Wincker, P. and Weissenbach, J.
TITLE Genomic exploration of the hemiascomycetous yeasts: 1. A set of
yeast species for molecular evolution studies
JOURNAL FEBS Lett. 487 (1), 3-12 (2000)
MEDLINE 20584711
REFERENCE 3 (bases 1 to 958)
Genoscope.
AUTHORS Direct Submission
TITLE Submitted (08-SEP-2000) Genoscope - Centre National de Sequencage,
JOURNAL 2 rue Gaston Cremieux, Cp 5706, 91057 EVRY cedex, FRANCE. (E-mail :
seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)

COMMENT

This GSS is part of a random genomic sequencing program of thirteen yeast species: Saccharomyces bayanus var. uvarum, Saccharomyces exiguus, Saccharomyces servazzii, Zygosaccharomyces rouxii, Saccharomyces kluyveri, Kluyveromyces thermotolerans, Kluyveromyces lactis var. lactis, Kluyveromyces marxianus var. marxianus, Pichia angusta, Debaryomyces hansenii var. hansenii, Pichia sorbitophila, Candida tropicalis and Yarrowia lipolytica. Genomic inserts of 3 to 5 kb were prepared and both extremities were sequenced. See keywords for description of this sequence and for the sequence of the other extremity of this insert.

FEATURES

source

1..958
Location/Qualifiers

/organism="Pichia farinosa"
/strain="CBS 7064"
/db_xref="taxon:4920"
/clone="AX0AA027B12"
/clone_lib="AX0AA"
/note="end : 17"

misc_feature

<36..>914
/note="similar to Saccharomyces cerevisiae ORF YOR381w [FRE3 ; strong similarity to ferric reductase Fre2p]"
/evidence=not_experimental
<36..>947

misc_feature

/note="similar to Saccharomyces cerevisiae ORF YNR060w [FRE4 ; strong similarity to Fre2p, and similarity to Fre1p]"
/evidence=not_experimental

misc_feature

<291..>914
/note="similar to Saccharomyces cerevisiae ORF YKL220c [FRE2 ; ferric (and cupric) reductase]
similar to Saccharomyces cerevisiae ORF YOR384w [FRE5 ; strong similarity to ferric reductase Fre2p]"
/evidence=not_experimental
BASE COUNT 237 a 222 c 246 g 226 t 27 others
ORIGIN

Query Match 2.4%; Score 50.6; DB 13; Length 958;
Best Local Similarity 43.7%; Pred. No. 0.0053;
Matches 219; Conservative 11; Mismatches 265; Indels 6; Gaps 1;

QY 884 actttcaacttttaccacaaatggtcagcatcgtctgcttcatgttagcgcgtcgtccat 943
DB 30 ACCATTATCGCTTCCACAGAGGGGGCGCAAGAGAGCTGTTTCATCATGGATTGATCCAC 89

QY 944 tcaatgttatgacgcgttcaggagttaaacgaggagtattccagtccttcttgtaagaa 1003
DB 90 GCCKGCGCYTCAGGCAACATTTGTTTCAGTACCATGACTACGCAAGCAAAATGCGCAA 149

QY 1004 ttctacttcagatgggaatagtagccacaattcttattgtccatcatcttttccagtc 1063
DB 150 ACCCAGTGGGTGGGTACCATTTGCCACGGGGCGGGGAATCATTTATGTGCCAAGGK 209

QY 1064 gagaaggtcttcaggaacccaggttatgaaattcttttacttattccaaagccatgaac 1123
DB 210 ATGCCCTTCTTGAGAAGGGCCTGGKACAGATTTSCCTTCTTATACACATTATCATGGCC 269

QY 1124 atcattgttatcatagctatgtattaccattgccacactaggatggatgggctggatc 1183
DB 270 GCCTTATTATTGGAGGAGCCTGGATCCCATGTGGAGGANCBKGGATACGTTTACMTCYAT 329

QY 1184 tggctcatgcttggaactcctctgcttcgcagaggttctgcgaattgtacgtatcatcag 1243
DB 330 MATGCMACBGTGCCCTCATGGGTCTATGGGTCTTTGANAGAGTGGGCGGCBTGGGGAATTCGAGTT 389

QY 1244 aacgaggtcttaagaccgccactttgtcgaccacagatgattcttaacgtatcaagatc 1303
DB 390 TTGTGATTCCCAAGTCGGAAGTAATAWMTBGCBGACGAGACGT-----TGAAGTG 443

QY 1304 tctgtcaagaagcctaagttcttcttcaagtggagcatttgccctatatgtacttt 1363
DB 444 GTCATTTCCAAAGCCTCTCTTACKGGAAGTCTATCCCGAGGAGACATGCCTTTATTCACTTC 503

Qy 1364 ctttcacaaatcagcctgg 1384
Db 504 TTGAGACCAACGCTGCTGG 524

RESULT 7
LOCUS CNS07902 851 bp DNA GSS 08-JUL-2001
DEFINITION T3 end of clone BB0AA031C09 of library BB0AA from strain CBS 4732 of *Pichia angusta*, genomic survey sequence.
ACCESSION AL435600
VERSION AL435600.1 GI:12219013
KEYWORDS GSS.
SOURCE *Pichia angusta*.
ORGANISM *Pichia angusta*.
REFERENCE Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes; Saccharomycetales; Saccharomycetaceae; *Pichia*.
AUTHORS 1 (bases 1 to 851)
Blandin,G., Llorente,B., Malpertuy,A., Wincker,P., Artiguenave,F. and Dujon,B.
TITLE Genomic exploration of the hemiascomycetous yeasts: 13. *Pichia angusta*
JOURNAL FEBS Lett. 487 (1), 76-81 (2000)
MEDLINE 20584723
REFERENCE 2 (bases 1 to 851)
AUTHORS Souciet,J.L., Aigle,M., Artiguenave,F., Blandin,G., Bolotin-Fukuhara,M., Bon,E., Brottier,P., Casaregola,S., de-Montigny,J., Dujon,B., Durrens,P., Lepingle,A., Llorente,B., Malpertuy,A., Neuveglise,C., Ozier-Kalogeropoulos,O., Potier,S., Saurin,W., Tekala,F., Toffano-Nioche,C., Wesolowski-Louvel,M., Wincker,P. and Weissenbach,J.
TITLE Genomic exploration of the hemiascomycetous yeasts: 1. A set of yeast species for molecular evolution studies
JOURNAL FEBS Lett. 487 (1), 3-12 (2000)
MEDLINE 20584711
REFERENCE 3 (bases 1 to 851)
AUTHORS Direct Submission
JOURNAL Submitted (08-SEP-2000) Genoscope - Centre National de Sequencage, 2 rue Gaston Cremieux, CP 5706, 91057 EVRY cedex, FRANCE. (E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)
COMMENT This GSS is part of a random genomic sequencing program of thirteen yeast species: Saccharomyces bayanus var. uvarum, Saccharomycetes exiguus, Saccharomycetes servazzii, Zygosaccharomyces rouxii, Saccharomycetes kluyveri, Kluyveromyces thermotolerans, Kluyveromyces lactis var. lactis, Kluyveromyces marxianus var. marxianus, *Pichia angusta*, Debaryomyces hansenii var. hansenii, *Pichia sorbitophila*, *Candida tropicalis* and *Yarrowia lipolytica*. Genomic inserts of 3 to 5 kb were prepared and both extremities were sequenced. See keywords for description of this sequence and for the sequence of the other extremity of this insert.
FEATURES
source Location/Qualifiers
1. .851
/organism="Pichia angusta"
/strain="CBS 4732"
/db_xref="taxon:4905"
/clone="BB0AA031C09"
/clone_lib="BB0AA"
/note="end : T3"
misc_feature <3..>822
/note="similar to Saccharomyces cerevisiae ORF YOR384w [FRES ; strong similarity to ferric reductase Fre2p]
1 putative frameshift(s)"
/evidence=not_experimental
BASE COUNT 204 a 196 c 201 g 246 t 4 others
ORIGIN

Query Match 2.3%; Score 48.2; DB 13; Length 851;
Best Local Similarity 50.8%; Pred. No. 0.024;
Matches 166; Conservative 2; Mismatches 150; Indels 9; Gaps 2;

Qy 1045 catcatcttcagtcgaagaggtcttcagaaacccaggtatataaaattcttact 1104

Wing, R.A. and Dean, R.A.
A BAC End Sequencing Framework to Sequence the Rice Genome
Unpublished (1998)
Contact: Wing RA
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Tel: 864 656 7288
Fax: 864 656 4293
Email: rwing@clemson.edu
Seq primer: TAATACGACTCACTATAGGG
Class: BAC ends
High quality sequence start: 32
High quality sequence stop: 406.
Location/Qualifiers
1. .805
/organism="Oryza sativa"
/strain="Japonica"
/cultivar="Nipponbare"
/db_xref="taxon:4530"
/clone="nbeb0002019f"
/clone_lib="CUGI Rice BAC Library (ECORI)"
/tissue_type="Leaf"
/lab_host="E. coli DH108"
/note="vector: pBACIndigo; Site_1: EcoRI; Site_2: EcoRI;
Rice is the most important food crop in the world. Half of
the world population, especially those inhabiting highly
populated areas of the humid tropics and subtropics, rely
on rice as their primary source of carbohydrate.
Monocotyledonous rice is a diploid plant (2n=24) with a
haploid genome equivalent of 431 Mbp (Arumuganathan and
Earle, 1991). The relatively small genome of rice, three
times larger than that of Arabidopsis, makes it suitable
for genomic studies. In order to facilitate positional
cloning, physical mapping and genome sequencing of rice,
we have constructed a BAC library from Oryza sativa, the
Nipponbare variety using EcoRI as the cloning enzyme. The
library contains 55,296 clones with an average insert size
of 121 Kb providing approximately 15 haploid genome
equivalents. The deep coverage allows the isolation a
particular sequence with a probability of 99.9 %. Three
high density filters, each containing 18,432 clones
(doubly spotted), represent the whole library for colony
screening and can be requested from the Clemson University
BAC/EST Resource Center (www.genome.clemson.edu)."

181 a 189 q 198 t

```

BASE COUNT      181 a      237 c      189 g      198 t
ORIGIN
Query Match      2.18; Score 44; DB 13; Length 805;
Best Local Similarity 47.2%; Pred. NO. 0.34;
Matches 134; Conservative 0; Mismatches 150; Indels 0; Gaps 0:

Qy 1801 tgggtcattcagtgaggatcaaaactcagatgagtcacctaagggtttccgatgacaagga 1860
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 606 TGTGTCACCCCACAAAGCAGCGCTGGCAGAGAGCTATGGCGGGCGTCTTGGCATGCA 547
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Qy 1861 agaactctgaatcaccgttagaatcccttaacaagggccagacctcaaaagcagctagttag 1920
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 546 TGCATAATAAACAAGAAATCCCATCTGTTTCAGAACTTGTAAACAACACCAAGCCCTTCCCAT 487
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Qy 1921 atcagagatcaaaattgtcagaactcgagaacacaacacatcattcttactcatcgcgacc 1980
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 486 GTAAGAAAACAGATCTTCTGTGACITGGGAATTAATTACACCGTTGGTCATCAATAATCCA 427
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Qy 1981 agcgacttcaatgacgactttaggaatgcagttgtgtacaaggtatcgattcttagctgaa 2040
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 426 ATGCACATAACAAGAGGTACTAGSAAAAAGGTATTAAGAAGGGTGGGCCAGGTATAAAG 367
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Qy 2041 gatgagtgtcgaactagagggaggagttttacttggtaagagc 2084
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 366 GTGAAGAGTTAAAAATGSPAATTGTACTTTTATAGTTGATTTGTGTC 323
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

RESULT	11
LOCUS	DEFINITION
ACCESSION	VERSION
KEYWORDS	SOURCE
ORGANISM	REFERENCE
AUTHORS	TITLE
JOURNAL	COMMENT

BASE COUNTRY ORIGIN	Query Match Best Loss Matches	
	Qy	847
	Db	81
	Qy	907
	Db	141
	Qy	967
	Db	201
	Qy	1027
	Db	261

Query Match	2.0%; Score 42.8; DB 13; Length 949;
Best Local Similarity	44.9%; Pred. No. 0.78;
Matches 151; Conservative 12; Mismatches 169; Indels 4; Gaps 1;	
QY 811	ttcttcccggtggttacatttttcggtatccggaacaccccttcattcccaatcacgg 870
DB 441	TTCTAATGTGCTTAAACGTTTCWATTTGGTGGACRTAACAACTTCCTATCTCGGCTACTAR 382
QY 871	atgagcttttagtactttccaacttttaccacaaatggctcagctacac-----gtctgctcca 926
DB 381	MTAGAAACAACACTACTWTTATACATACACCAAGGATTCGCATACAGTAGTGTGCTATATCA 322
QY 927	tgttagccgctgcctccattcaatcggtatgaacccgcttcagaggttaaacagaggagattacc 986
DB 321	TAGTTATTTGTTTACATTCCTATCGCTATRTTATCTATGCTGTATTTGRCAGACGCCCTTTT 262
QY 987	agtccttgaagaaatttactctcagatggggaatagtagccacaattcttatgtcca 1046
DB 261	CCTACTGGATTCTACAGGACTGCTACAGATGGGTTGTTGCTACTGTGCTGGTACAA 202
QY 1047	tcatcattttccagtcggagaaggtcttcaggaaaccgaggttatgaattcttctactta 1106
DB 201	TCATATTTTATTCAGGTATATCTGCTGCTCCTGCTTAACTACTAGACTTTCTGTGCA 142
QY 1107	ttcacaagccatgaacatcatgtttatcatagcta 1142
DB 141	TCCACATTGTGATRGACGATTTCTTTTAGTCGGTA 106
RESULT 13	
CNS0039G	
LOCUS	CNS0039G 1101 bp DNA GSS 03-JUN-1999
DEFINITION	Drosophila melanogaster genome survey sequence TET3 end of BAC #
	BACR08K10 of RPCI-98 library from Drosophila melanogaster (fruit
	fly), genomic survey sequence.
ACCESSION	AL063921
VERSION	AL063921.1 GI:4941778
KEYWORDS	GSS.
SOURCE	fruit fly.
ORGANISM	Drosophila melanogaster
	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
	Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
	Muscomorpha; Phylloidea; Phylloidea; Drosophila;
	1 (bases 1 to 1101)
REFERENCE	Genoscope.
AUTHORS	Direct Submission
TITLE	Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
JOURNAL	BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
	- Web : www.genoscope.cns.fr)
COMMENT	Determination of this BAC end sequence was carried out as part of a
	collaboration with the Berkeley Drosophila Genome Project (BDGP).
	The BDGP is constructing a physical map of the Drosophila
	melanogaster genome using these BACs. For further information
	please see http://www.fruitfly.org The BDGP Drosophila
	melanogaster BAC library was prepared by Kazutoyo Osogawa and
	Aaron Mammeter at the Roswell Park Cancer Institute in Buffalo,
	NY. The library is named RPCI-98 and was constructed by partial
	EcoRI digestion of Drosophila DNA provided by the BDGP from the
	isogenic strain y; cn bw sp, the same strain used for the BDGP's
	p1 and EST libraries. A more detailed description of the library
	and how to order individual BAC clones, the entire library, or
	filters for hybridization from the BACPAC Resource Center can be
	found at http://bacpac.med.buffalo.edu/drosophila_bac.htm .
FEATURES	Location/Qualifiers
source	1..1101
	/organism="Drosophila melanogaster"
	/db_xref="taxon:7227"
	/clone_lib="RPCI-98"
	/clone="BACR08K10"
	/notes="end : 1101"

```
BASE COUNT      201 a      64 c      131 g      202 t      503 others
ORIGIN

Query Match      2.0%; Score 42.8; DB 13; Length 1101;
Best Local Similarity 17.3%; Pred. No. 0.82;
Matches 109; Conservative 252; Mismatches 264; Indels 5; Gaps 2;

QY 545 acattatggaacattctgttaagaagctcttattaccacaaagcgtttacaaagactac 604
Db 416 AWAATAAATAATTTWAAWAAWAAATAATTTWAAWAAWAAWAAWAAWAAWAAWAAWAA 475

QY 605 aacgagagaaacttctatcttctggaacgcttgccttccattcaactttcaacactcaggcaaa 664
Db 476 AAAAAAATTTTWTWTWTWTWTWTWTWTWTWTWTWTWTWTWTWTWTWTWTWTWTWTWT 535

QY 665 ggaactgtagtcttattcttctggaacgcttgccttccattcaactttcaacactcaggcaaa 724
Db 536 AWATAAATTTTWTWTWTWTWTWTWTWTWTWTWTWTWTWTWTWTWTWTWTWTWTWTWT 595

QY 725 atcaagtgcacacattctacgatacctagatgagagagagagagagagagagagagagagag 784
Db 596 TWTWAAWAAWAAWAAWAAWAAWAAWAAWAAWAAWAAWAAWAAWAAWAAWAAWAAWAAW 655

QY 785 cgcgtgactgactgagcgaacgctcttcttcccggtggtgtacatttccggtatccgg 844
Db 656 HAHAAWAAWAAWAAWAAWAAWAAWAAWAAWAAWAAWAAWAAWAAWAAWAAWAAWAAW 715

QY 845 acaaccccttaccacacacacacacacacacacacacacacacacacacacacacacacac 902
Db 716 HHWYAHYMMYMMYMMYMMYMMYMMYMMYMMYMMYMMYMMYMMYMMYMMYMMYMMYMM 775

QY 903 aatggcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 962
Db 776 AAWAAWAAWAAWAAWAAWAAWAAWAAWAAWAAWAAWAAWAAWAAWAAWAAWAAWAAW 835

QY 963 caggagtaaac---gaggagatattccagctctcttgaagaataattctactcagatcg 1019
Db 836 WTHWYHHTWTHWTHWTHWTHWTHWTHWTHWTHWTHWTHWTHWTHWTHWTHWTHWTHW 895

QY 1020 gaatagtagcacacattctatgtccatccatccatccatccatccatccatccatccatcc 1079
Db 896 ACHAAHHTHMCCHHHHMCCHHHHHTWTHWTHWTHWTHWTHWTHWTHWTHWTHWTHWTHW 955

QY 1080 accgaggtatgaattcttcttcttcttcttcttcttcttcttcttcttcttcttctatag 1139
Db 956 CCMHHHCHHCHHCHHCHHCHHCHHCHHCHHCHHCHHCHHCHHCHHCHHCHHCHHCHH 1015

1140 ctatgtattaccattgccacacacacacacacacacacacacacacacacacacacacacac 1169
: | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1016 HAMTTATWMMHWHWAAWAAWAAWAAWAAWAAWAAWAAWAAWAAWAAWAAWAAWAAWAA 1045

RESULT 14
CNS07DHO      1043 bp      DNA      GSS      08-JUL-2001
LOCUS
DEFINITION
T3 end of clone BD0AA012G01 of library BD0AA from strain CBS 94 of
Candida tropicalis, genomic survey sequence.
ACCESSION
AL440482
VERSION
AL440482.1 GI:12223893
KEYWORDS
GSS.
SOURCE
Candida tropicalis.
ORGANISM
Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
Saccharomycetales; microsporid Saccharomycetales; Candida.
REFERENCE
1 (bases 1 to 1043)
AUTHORS
Blandin,G., Ozler-Kalogeropoulos,O., Wincker,P., Artiguenave,F. and
Dujon,B.
TITLE
Genomic exploration of the hemiascomycetous yeasts: 16. Candida
tropicalis
JOURNAL
FEBS Lett. 487 (1), 91-94 (2000)
MEDLINE
20584726
REFERENCE
2 (bases 1 to 1043)
AUTHORS
Souciet,J.L., Aigle,M., Artiguenave,F., Blandin,G.,
Bolotin-Fukuhara,M., Bon,E., Brottier,P., Casaregola,S.,
de-Montigny,J., Dujon,B., Durand,P., Lepingle,A., Llorente,B.,
Malpertuy,A., Neuveglise,C., Ozler-Kalogeropoulos,O., Potier,S.,
Saurin,W., Tekia,F., Toffano-Nioche,C., Wesolowski-Louvel,M.,
Wincker,P. and Weissenbach,J.
TITLE
Genomic exploration of the hemiascomycetous yeasts: 1. A set of
yeast species for molecular evolution studies
JOURNAL
FEBS Lett. 487 (1), 3-12 (2000)
MEDLINE
20584711
REFERENCE
3 (bases 1 to 1043)
AUTHORS
Genoscope.
TITLE
Direct Submission
JOURNAL
Submitted (08-SEP-2000) Genoscope - Centre National de Sequencage,
2 rue Gaston Crémieux, CP 5706, 91057 EVRY cedex, FRANCE. (E-mail :
seque@genoscope.cns.fr - Web :
http://www.genoscope.cns.fr)
COMMENT
This GSS is part of a random genomic sequencing program of thirteen
yeast species: Saccharomyces bayanus var. uvarum, Saccharomyces
exiguus, Saccharomyces servazzii, Zygosaccharomyces rouxii,
Saccharomyces kluyveri, Kluyveromyces thermotolerans, Kluyveromyces
lactis var. lactis, Kluyveromyces marxianus var. marxianus, Pichia
angusta, Debaryomyces hansenii var. hansenii, Pichia sorbitophila,
Candida tropicalis and varrowia lipolytica. Genomic inserts of 3 to
5 kb were prepared and both extremities were sequenced. See
keywords for description of this sequence and for the sequence of
the other extremity of this insert.
FEATURES
Location/Qualifiers
1..1043
source
/organism="Candida tropicalis"
/strain="CBS 94"
/db_xref="taxon:5482"
/clone_lib="BD0AA012G01"
/clone="BD0AA"
/notes="end : T3"
misc_feature
<2..>985
/note="similar to Saccharomyces cerevisiae ORF YKL220c [
FRE2; ferric (and cupric) reductase ]
similar to Saccharomyces cerevisiae ORF YNR060w [ FRE4 ;
strong similarity to Fre2p, and similarity to Fre1p ]
similar to Saccharomyces cerevisiae ORF YOR381w [ FRE3 ;
strong similarity to ferric reductase Fre2p ]
similar to Saccharomyces cerevisiae ORF YOR384w [ FRE5 ;
strong similarity to ferric reductase Fre2p ]"
/evidence="not_experimental"
BASE COUNT      287 a      173 c      235 g      345 t
ORIGIN

Query Match      2.0%; Score 41; DB 13; Length 1043;
Best Local Similarity 52.0%; Pred. No. 2.5;
Matches 92; Conservative 0; Mismatches 85; Indels 0; Gaps 0;

QY 1072 cttcagaacccaggttatgaattcttcttcttcttcttcttcttcttcttcttcttcttctt 1131
Db 1 CTTCAGAAGACGATGGTATGAATGCTTTTCTTCTATACACATCGTTATGCTGCCATTTA 60

QY 1132 tatcatgactatgtattaccattgccacacacacacacacacacacacacacacacacacacac 1191
Db 61 TATTCTGGGTACGTGGATCCATGTTGATGATCTTGGATATGTTTGGTCTCTGTACGCCAG 120

QY 1192 ggcgtgcatctctgcttcgcagaggttctgcgcgaattgtacgtatcatcatcagaacgg 1248
Db 121 TTTTGCACCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 177

RESULT 15
BG714799
LOCUS
DEFINITION
BG714799 2149 bp mRNA EST 08-MAY-2001
602677117F1 NIH_MGC_96 Homo sapiens cDNA clone IMAGE:4799501 5',
mRNA sequence.
ACCESSION
BG714799
VERSION
BG714799.1 GI:13993730
KEYWORDS
EST.
```


